Supplementary Materials
to "Structural diversity and conservation of PilZ-related domains"
by M. Y. Galperin and S.-H. Chou

Table S1. PilZ-related domains with known 3D structure

Organism, locus tag	Protein name, UniProt entry	Domains in structure	PDB entry, resolution	Bound c-di-GMP	Reference
Pseudomonas aeruginosa PA4608	PilZ or MapZ, CDGBP_PSEAE	PilZ only	1YWU (NMR) 2L74 (NMR) 5XLY_B, 1.76 Å 5Y4R_D, 2.30 Å	No Yes Yes Yes	(1-3)
Xanthomonas citri Xac1133	PilZ, Q8PND9_XANAC	PilZ only	3CNR_A, 1.90 Å 4FOU_C, 2.10 Å	No No	(4, 5)
Xanthomonas campestris Xcc1024	PilZ, Xc1028, Q8PBU4_XANCP	PilZ only	3DSG_A, 2.09 Å 4F48_B, 3.00 Å	No No	(6, 7)
Legionella pneumophila lpg0364	PilZ, Q5ZYK8_LEGPH	PilZ only	4Q63_A, 1.95 Å	No	N/A
Rhodobacter sphaeroides RSP_0333	BcsA, Q3J125_RHOS4	Glyco_trans_2 - PilZ	4HG6_A, 3.25 Å 4P00_A, 3.25 Å 4P02_A, 2.65 Å 5EIY_A, 2.95 Å 5EJ1_A, 3.40 Å 5EJZ_A, 2.95 Å	No Yes Yes Yes Yes Yes	(8-10)
Komagataeibacter xylinus AcsAB	CelA, P0CW87, ACSA1_KOMXY	PilZ only	4I86, 2.10 Å	No	(11)
Pseudomonas aeruginosa PA3542	Alg44, Q9HY69 ALG44_PSEAE	PilZ only	4RT0, 1.8 Å 4RT1, 1.7 Å 4XRN, 2.0 Å	Yes Yes Yes	(12)
Escherichia coli YcgR (b1194)	YcgR, P76010, YCGR_ECOLI	PilZN-PilZ	5Y6F, 2.30 Å	Yes	(13)
		PilZN (YcgR)	5Y6H, 1.77 Å	No	
		PilZ	5Y6G, 2.30 Å	Yes	
Pseudomonas putida PP_4397	FlgZ, Q88EQ6, YCGR_PSEPK	PilZN-PilZ (YcgR-PilZ)	2GJG, 2.25 Å 3KYF, 2.10 Å	No Yes	(14)
Vibrio cholerae VCA0042 (=VC0395_0091)	PlzD, Q9KNC3 YCGRL_VIBCH	PilZNR-PilZ (YcgR_2-PilZ)	1YLN, 2.20 Å 2RDE, 1.92 Å 3KYG, 2.10 Å	No Yes Yes	(14, 15)
Bacillus subtilis BSU22910	MotI or DgrA, YPFA_BACSU	PilZNR-PilZ (YcgR_2-PilZ)	5VX6, 3.2 Å	Yes	(16)
Klebsiella pneumoniae KPK_0839	MrkH, G3FT00_KLEPN	PilZN-PilZ (YcgR-PilZ)	5EJL, 2.30 Å 5KEC, 1.95 Å 5KED, 2.65 Å 5KGO, 2.90 Å	Yes No No Yes	(17, 18)
Xanthomonas campestris XccB100_2234	aPilZ or Xcc0612, B0RT03_XANCB	PilZ_2	3PH1, 2.1 Å 3RQA, 2.1 Å	No No	(19)
Geobacter sulfurreducens GSU3033	Q748G9_GEOSL	DUF5634_N	2L1T (NMR)	No	N/A

Table S2. Structural similarity of PilZ domains according to Dali and VAST

Organism, protein name (genomic locus tag)[a]	**PDB entry**	**Alignment length, % identity, RMSD**[b]	
		Dali	**VAST**
Pseudomonas aeruginosa PilZ, MapZ (PA4608)	2L74_A	105 aa, 100%, 2.1 Å	102 aa, 100%, 1.5 Å
Xanthomonas citri PilZ (Xac1133)	3CNR_B	72 aa, 13%, 2.9 Å	64 aa, 12%, 2.6 Å
Xanthomonas campestris PilZ (Xcc1024)	3DSG_B	72 aa, 13%, 3.0 Å	61 aa, 16%, 2.4 Å
Legionella pneumophila PilZ (Lpg0364)	4Q63_A	88 aa, 11%, 2.7 Å	78 aa, 10%, 2.1 Å
Rhodobacter sphaeroides BcsA (RSP_0333) PilZ domain, aa 578-691	5EJZ_A	98 aa, 12%, 2.6 Å	72 aa, 15%, 1.7 Å
Komagataeibacter xylinus AcsAB, PilZ domain	4I86_B	92 aa, 20%, 2.6 Å	91 aa, 20%, 2.5 Å
Pseudomonas aeruginosa Alg44 (PA3542), PilZ domain	4RT1_B	101 aa, 17%, 2.6 Å	98 aa, 17%, 2.2 Å
Escherichia coli YcgR (b1194)			
PilZ domain, aa 111-244	5Y6G_A	101 aa, 17%, 2.7 Å	83 aa, 19%, 2.2 Å
YcgR domain, aa 1-110	5Y6H_A	80 aa, 4%, 2.9 Å	59 aa, 2%, 1.7 Å
Pseudomonas putida FlgZ (PP_4397) PilZ domain, aa 121-238	3KYF_A	101 aa, 12%, 2.5 Å	93 aa, 13%, 2.3 Å
Klebsiella pneumoniae MrkH			
PilZ domain, aa 107-236	5KEC_D	104 aa, 11%, 3.1 Å	87 aa, 11%, 2.2 Å
MrkH_N domain, aa 1-106	5KEC_D	77 aa, 10%, 2.9 Å	65 aa, 11%, 1.8 Å
Vibrio cholerae PlzD (VCA0042)			
PilZ domain, aa 134-247	2RDE_A	99 aa, 14%, 3.1 Å	89 aa, 17%, 2.2 Å
YcgR_2 domain, aa 24-130	2RDE_A	75 aa, 9%, 2.2 Å	66 aa, 11%, 1.6 Å
Bacillus subtilis MotI (BSU22910)			
PilZ domain, aa 97-208	5VX6_A	104 aa, 16%, 3.2 Å	96 aa, 17%, 1.9 Å
MotI_N domain, aa 1-96	5VX6_A		69 aa, 7%, 2.1 Å
Xanthomonas campestris Xcc0612 (XccB100_2234)	3RQA_A	93 aa, 12%, 2.9 Å	84 aa, 13%, 2.2 Å
Geobacter sulfurreducens GSU3033 (aa 1-102)	2L1T_A	85 aa, 5%, 3.1 Å	65 aa, 5%, 2.2 Å

[a] – For multidomain structures, the amino acid (aa) boundaries of the canonical PilZ domains are shown in parentheses.

[b] – The overlap length, percent identity, and RMSD values were taken from DALI (20) and VAST (21) alignments of the respective structures against the stand-alone PilZ domain in the *P. aeruginosa* protein MapZ (PDB: 5XLY_B).

Table S3. Examples of tetrameric PilZ domains

Organism, taxonomy	GenBank accession	UniProt entry	c-di-GMP motifs[a]
Bacteroidetes			
Sphingobacteriales bacterium	RYE43434	A0A4Q3G9X4	Yes
Deferribacteres			
Calditerrivibrio nitroreducens	ADR18856	E4THN1	Yes
Deferribacter desulfuricans	BAI79927	D3PBF4	Yes
Geovibrio thiophilus	QAR34335	A0A3R5Z107	No
Denitrovibrio acetiphilus	ADD69612	D4H6D8	No
Seleniivibrio woodruffii	TCK61742	A0A4R1KBA3	No
Nitrospinae			
Nitrospinae bacterium	HBK81472	A0A355B8P2	Yes
Nitrospinae bacterium	MAE05034	A0A2D5UAS9	Yes
Nitrospirae			
Leptospirillum ferriphilum ML-04	AFS54475	J9ZCV6	No
Ca. Magnetobacterium bavaricum	KJU82198	A0A0F3GNC4	Yes
Nitrospira defluvii	CBK42057	D8PFM0	Yes
Nitrospira moscoviensis NSP M-1	ALA58562	A0A0K2GD75	Yes
Nitrospira lenta BS10	SPP65144	A0A330L5V6	Yes
Ca. Nitrospira inopinata	CUQ67246	A0A0S4KVQ2	Yes
Ca. Nitrospira nitrificans COMA2	CUS31512	A0A0S4L4J9	Yes
Thermodesulfovibrio aggregans	GAQ94083	A0A0U9HLW8	Yes
Thermodesulfovibrio yellowstonii DSM 1134	ACI20485	B5YK90	Yes
Planctomycetes			
Ca. Scalindua brodae	KHE92567	A0A0B0EHM4	Yes
Ca. Scalindua japonica	GAX60402	A0A286TWZ4	Yes
Candidatus Tectomicrobia			
Ca. Tectomicrobia bacterium	OGL68512	A0A1F7TR77	Yes
Ca. Tectomicrobia bacterium	OGL59976	A0A1F7T1V7	Yes
Thermodesulfobacteria			
Caldimicrobium thiodismutans TF1	BAU22708	A0A0U5AKT9	Yes
Thermodesulfatator indicus DSM 15286	AEH44194	F8AE58	Yes
Thermodesulfobacterium commune	AIH03543	A0A075WT12	Yes
Thermodesulfobacterium geofontis	AEH23615	F8C3H8	Yes
Thermosulfurimonas dismutans	OAQ21198	A0A179D6G4	Yes
Proteobacteria			
Acidithiobacillia			
Acidithiobacillus caldus SM-1	AEK58128	F9ZN85	No
Acidithiobacillus thiooxidans	OCX69789	A0A1C2I1C8	No
Alphaproteobacteria			
Magnetococcus marinus MC-1	ABK44455	A0L912	Yes
	ABK44456	A0L913	No
Magnetofaba australis IT-1	OSM02482	A0A1Y2K3E3	Yes
	OSM02483	A0A1Y2K3B3	No

Betaproteobacteria			
Burkholderiales			
Chitinimonas taiwanensis DSM 18899	SFZ79374	A0A1K2HRK9	No
Janthinobacterium sp. B9-8	AMC34180	A0A0X8GL31	No
Thiomonas arsenitoxydans	CAZ88219	D6CQD9	No
Thiomonas intermedia K12	ADG30598	D5X0F2	No
Neisseriales			
Andreprevotia lacus DSM 23236	SMC24803	A0A1W1XLH4	No
Aquitalea magnusonii SM6	KJV27573	A0A0F3K8N4	No
Chromobacterium haemolyticum H4137	OQS41564	A0A1W0D3J5	No
Chromobacterium violaceum ATCC 12472	AAQ61178	Q7NSA9	No
Iodobacter sp. BJB302	PHV03655	A0A2G3JGH2	No
Vogesella sp. EB	KMJ54682	A0A0J5GJX3	No
Nitrosomonadales			
Methylobacillus sp. MM3	OAJ71600	A0A1A9NQ49	No
Novimethylophilus kurashikiensis La2-4	GBG13164	A0A2R5F957	No
Sulfuricella denitrificans skB26	BAN35494	S6ALD5	No
Sulfurirhabdus autotrophica DSM 100309	TCV90146	A0A4R3YDM4	No
Sulfuritortus calidifontis DSM 103923	TCS70553	A0A4R3JTB5	No
Rhodocyclales			
Rhodocyclales bacterium GWA2_65_20	OHC67057	A0A1G3HE94	No
Gammaproteobacteria			
Alteromonadales			
Alteromonas lipolytica	OFI36423	A0A1E8FLX1	No
Alteromonas macleodii	AFT74208	A0A1E7DF04	No
Pseudoalteromonas atlantica T6c	ABG40247	Q15V41	No
Cellvibrionales			
Cellvibrio japonicus Ueda107	ACE83634	B3PFR1	Yes
Congregibacter litoralis KT71	EAQ98448	A4A601	No
Ca. Endobugula sertula	ODS24617	A0A1D2QSR9	Yes
Marinimicrobium koreense DSM 16974	ROQ20302	A0A3N1NN93	Yes
Simiduia agarivorans SA1 = DSM 21679	AFU99140	K4KLP2	Yes
Saccharophagus degradans 2-40	ABD81068	Q21JR1	Yes
Teredinibacter turnerae T7901	ACR11340	C5BU89	Yes
Chromatiales			
Acidihalobacter prosperus	AOV16626	A0A1D8K6R8	No
Ectothiorhodospira haloalkaliphila	AHK79421	W8KV09	No
Ectothiorhodospira sp. PHS-1	EHQ51976	H1G2D9	No
Imhoffiella purpurea	EXJ16754	W9VBD1	Yes
Marichromatium purpuratum 984	AHF03627	W0DYU7	No
Nitrosococcus halophilus Nc 4	ADE15429	D5BV79	No
Thioalkalivibrio sp. K90mix	ADC71663	D3S9F8	No
Thioalkalivibrio sulfidiphilus HL-EbGr7	ACL73055	B8GT40	No
Thiocystis violascens DSM 198	AFL72424	I3Y606	No
Thiorhodococcus drewsii AZ1	EGV30492	G2E330	No
Thiorhodospira sibirica ATCC 700588	EGZ47185	G4E4N4	No
Immundisolibacterales			
Immundisolibacter cernigliae	ANX04033	A0A1B1YTP7	No
Legionellales			
Fluoribacter bozemanae WIGA	KTC74835	A0A0W0RUT1	Yes
Legionella pneumophila	AAU27996	Q5ZU81	Yes
Tatlockia micdadei NZ2015	ARG97104	A0A098GHK6	Yes

Methylococcales			
Methylobacter tundripaludum SV96	EGW20922	G3J150	Yes
Methylocaldum marinum	BBA37268	A0A250L021	Yes[b]
Methylomagnum ishizawai	SMF95440	A0A1Y6CYY6	Yes
Methylomonas methanica MC09	AEG00094	G0A1N3	Yes
	AEG00685	F9ZV75	Yes
Nevskiales			
Fontimonas thermophila	SFF33484	A0A1I2HTV9	No
Hydrocarboniphaga daqingensis	SHG49207	A0A1M5K8U8	No
Hydrocarboniphaga effusa AP103	EIT69541	I8I1H1	No
Nevskiales bacterium	MBV61603	A0A2E9ZY24	No
Solimonas aquatica	SEQ62310	A0A1H9HIZ0	No
Steroidobacter denitrificans	AMN46591	A0A127FAD2	No
Pseudomonadales			
Acinetobacter baumannii	SSU13771	A0A1G5DWA1	Yes[b]
Pseudomonas aeruginosa PAO1	AAG06377	Q9HZL5	Yes[b]
Pseudomonas aeruginosa UCBPP-PA14	ABJ12226	A0A0H2ZBV1	Yes[b]
Pseudomonas fluorescens Pf0-1	ABA75597	Q3K9F7	No
Pseudomonas putida KT2440	AAN67766	Q88KY6	No
Pseudomonas stutzeri A1501	ABP80294	A4VMU3	No
Pseudomonas syringae pv. syringae B728a	AAY36947	Q4ZV75	Yes?[c]
Thiopseudomonas denitrificans DSM 28679	TDQ40262	A0A4R6U470	Yes?[c]
Salinisphaerales			
Salinisphaeraceae bacterium CPC72	MAA76332	A0A2D4SGP5	No
Sedimenticola			
Sedimenticola thiotaurini	AKH19540	A0A0F7JXL1	No
Thiotrichales			
Cycloclasticus sp. strain P1	AFT66357	K0C3W4	Yes
Hydrogenovibrio crunogenus XCL-2	ABB42072	Q31FK1	Yes
Methylophaga frappieri JAM7	AFJ02461	I1YHR8	Yes
Thiomicrospira aerophila AL3	AHF02343	W0DYJ5	Yes
Vibrionales			
Vibrio cholerae O1 biovar El Tor str. N16961	AAF95033	Q9KQW3	No
Aliivibrio fischeri ES114	AAW87531	Q5E0B5	No
Enterovibrio norvegicus DSM 15893	SFP54731	A0A1I5R8T8	No
Photobacterium profundum SS9	CAG20776	Q6LPK0	No
Vibrio vulnificus YJ016	BAC95128	Q7MIZ9	No
Xanthomonadales			
Aquimonas voraii	SDD89738	A0A1G6YH90	No
Dyella japonica DSM 16301	KLD62119	A0A0G9GYY7	No
Dyella thiooxydans	AND67566	A0A160MWT4	No
Frateuria aurantia DSM 6220	AFC86459	H8L3D7	No
Luteibacter rhizovicinus DSM 16549	APG02804	A0A1L3EP40	No
Luteibacter yeojuensis	KJV36288	A0A0F3KYJ5	No
Luteimonas sp. FCS-9	KLJ02740	A0A0H1AUJ2	No
Lysobacter defluvii IMMIB APB-9	KGO98734	A0A0A0M8Z6	No
Lysobacter spongiicola DSM 21749	SKA21588	A0A1T4S0W8	No
Mizugakiibacter sediminis	GAP65255	A0A0K8QJZ1	No
Pseudoxanthomonas dokdonensis	KRG71076	A0A0R0CZ69	No
Pseudoxanthomonas spadix BD-a59	AER55289	G7UMX8	No
Rhodanobacter sp. RIFOXYA1_FULL_67_6	OHC44793	A0A1G3FKY0	No
Stenotrophomonas maltophilia R551-3	ACF51593	B4SIK4	No

Deltaproteobacteria			
Desulfarculus baarsii DSM 2075	ADK85379	E1QLB4	Yes
Desulfatibacillum alkenivorans DSM 16219	SHL13587	A0A1M6Y5T2	Yes
Desulfonatronospira thiodismutans ASO3-1	EFI34632	D6SPF6	Yes?[c]
Desulfovibrio alaskensis G20	ABB38944	Q30ZF2	Yes?[c]
Desulfovibrio magneticus RS-1	BAH75713	C4XSL6	
Desulfurivibrio alkaliphilus AHT 2	ADH85899	D6Z2X4	Yes
Pseudodesulfovibrio aespoeensis Aspo-2	ADU62726	E6VYG3	Yes?[c]
Syntrophus aciditrophicus SB	ABC77079	Q2LSL7	Yes
Syntrophus gentianae DSM 8423	SEM26456	A0A1H7WXW1	Yes
Zetaproteobacteria			
Mariprofundus ferrooxydans PV-1	EAU54705	Q0EZJ9	Yes
Zetaproteobacteria bacterium CG2_30_59_37	OIQ01522	A0A1J5IWJ7	Yes

[a] – Presence of RxxxR and [D/N]hSxxG motifs, which is an indication of the ability to bind c-di-GMP.

[b] – The first residue of the [D/N]hSxxG motif is Ser, which still allowed c-di-GMP binding in *P. aeruginosa* PA2989.

[c] – A modified [D/N]hSxxG motif, c-di-GMP binding has not been tested.

References

1. **Ramelot TA, Yee A, Cort JR, Semesi A, Arrowsmith CH, Kennedy MA.** 2007. NMR structure and binding studies confirm that PA4608 from *Pseudomonas aeruginosa* is a PilZ domain and a c-di-GMP binding protein. Proteins **66:**266-271.
2. **Zhu Y, Yuan Z, Gu L.** 2017. Structural basis for the regulation of chemotaxis by MapZ in the presence of c-di-GMP. Acta Crystallogr D Struct Biol **73:**683-691.
3. **Yan XF, Xin L, Yen JT, Zeng Y, Jin S, Cheang QW, Fong R, Chiam KH, Liang ZX, Gao YG.** 2018. Structural analyses unravel the molecular mechanism of cyclic di-GMP regulation of bacterial chemotaxis via a PilZ adaptor protein. J Biol Chem **293:**100-111.
4. **Guzzo CR, Salinas RK, Andrade MO, Farah CS.** 2009. PILZ protein structure and interactions with PILB and the FIMX EAL domain: implications for control of type IV pilus biogenesis. J Mol Biol **393:**848-866.
5. **Guzzo CR, Dunger G, Salinas RK, Farah CS.** 2013. Structure of the PilZ-FimXEAL-c-di-GMP complex responsible for the regulation of bacterial type IV pilus biogenesis. J Mol Biol **425:**2174-2197.
6. **Li TN, Chin KH, Liu JH, Wang AH, Chou SH.** 2009. XC1028 from *Xanthomonas campestris* adopts a PilZ domain-like structure without a c-di-GMP switch. Proteins **75:**282-288.
7. **Chin KH, Kuo WT, Yu YJ, Liao YT, Yang MT, Chou SH.** 2012. Structural polymorphism of c-di-GMP bound to an EAL domain and in complex with a type II PilZ-domain protein. Acta Crystallogr D Biol Crystallogr **68:**1380-1392.
8. **Morgan JL, Strumillo J, Zimmer J.** 2013. Crystallographic snapshot of cellulose synthesis and membrane translocation. Nature **493:**181-186.
9. **Morgan JL, McNamara JT, Zimmer J.** 2014. Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. Nat Struct Mol Biol **21:**489-496.
10. **Morgan JL, McNamara JT, Fischer M, Rich J, Chen HM, Withers SG, Zimmer J.** 2016. Observing cellulose biosynthesis and membrane translocation in crystallo. Nature **531:**329-334.

11. **Fujiwara T, Komoda K, Sakurai N, Tajima K, Tanaka I, Yao M.** 2013. The c-di-GMP recognition mechanism of the PilZ domain of bacterial cellulose synthase subunit A. Biochem Biophys Res Commun **431:**802-807.

12. **Whitney JC, Hay ID, Li C, Eckford PD, Robinson H, Amaya MF, Wood LF, Ohman DE, Bear CE, Rehm BH, Howell PL.** 2011. Structural basis for alginate secretion across the bacterial outer membrane. Proc Natl Acad Sci USA **108:**13083-13088.

13. **Hou Y, Li D-F, Wang D-C.** 2013. Crystallization and preliminary X-ray analysis of the flagellar motor `brake' molecule YcgR with c-di-GMP from *Escherichia coli*. Acta Crystallogr F Struct Biol Cryst Commun **69:**663-665.

14. **Ko J, Ryu KS, Kim H, Shin JS, Lee JO, Cheong C, Choi BS.** 2010. Structure of PP4397 reveals the molecular basis for different c-di-GMP binding modes by Pilz domain proteins. J Mol Biol **398:**97-110.

15. **Benach J, Swaminathan SS, Tamayo R, Handelman SK, Folta-Stogniew E, Ramos JE, Forouhar F, Neely H, Seetharaman J, Camilli A, Hunt JF.** 2007. The structural basis of cyclic diguanylate signal transduction by PilZ domains. EMBO J **26:**5153-5166.

16. **Subramanian S, Gao X, Dann CE, 3rd, Kearns DB.** 2017. MotI (DgrA) acts as a molecular clutch on the flagellar stator protein MotA in *Bacillus subtilis*. Proc Natl Acad Sci USA **114:**13537-13542.

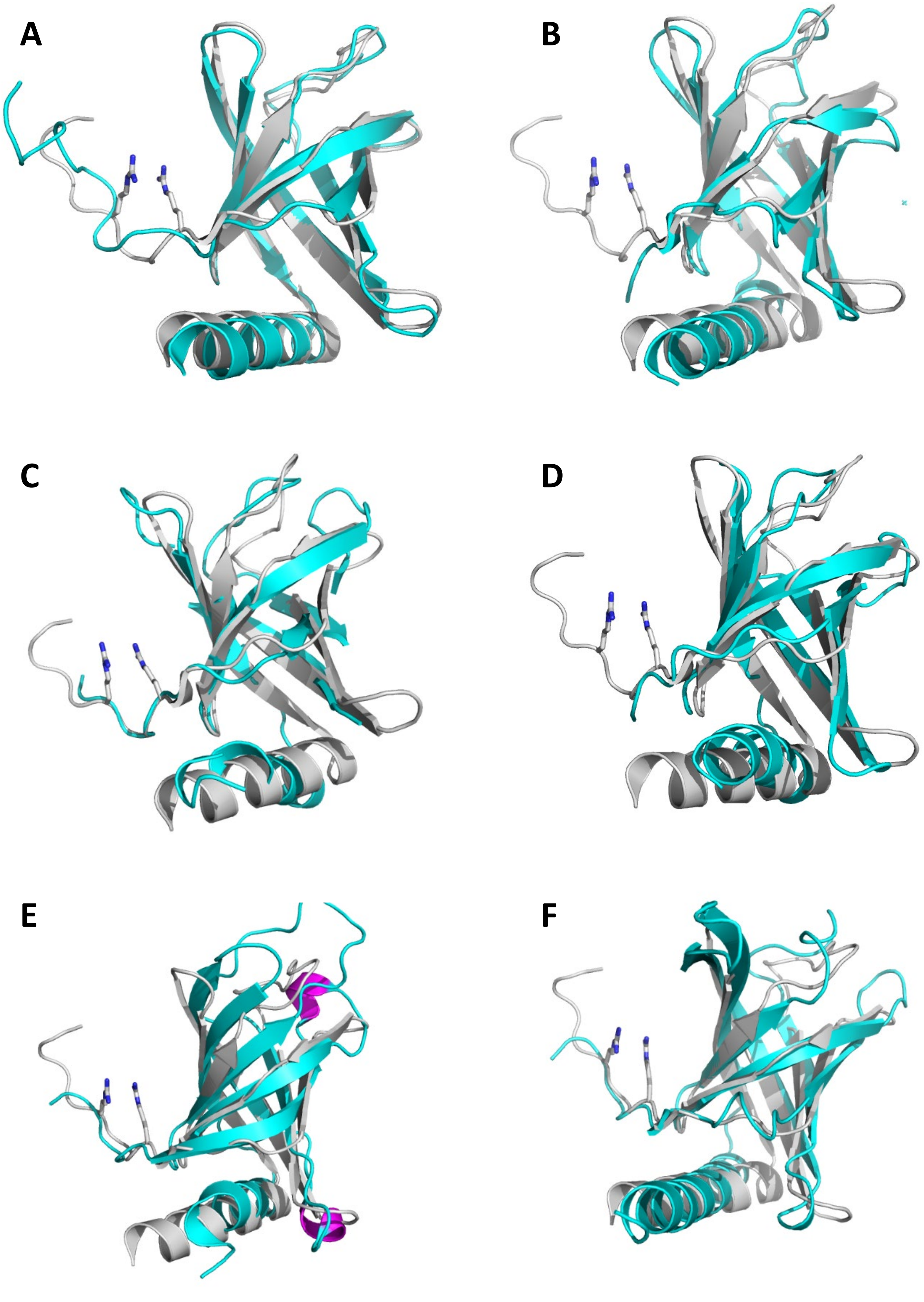
17. **Schumacher MA, Zeng W.** 2016. Structures of the activator of *K. pneumonia* biofilm formation, MrkH, indicates PilZ domains involved in c-di-GMP and DNA binding. Proc Natl Acad Sci USA **113:**10067-10072.

18. **Wang F, He Q, Su K, Gao F, Huang Y, Lin Z, Zhu D, Gu L.** 2016. The PilZ domain of MrkH represents a novel DNA binding motif. Protein Cell **7:**766-772.

19. **Li TN, Chin KH, Fung KM, Yang MT, Wang AH, Chou SH.** 2011. A novel tetrameric PilZ domain structure from xanthomonads. PLoS One **6:**e22036.

20. **Holm L, Laakso LM.** 2016. Dali server update. Nucleic Acids Res **44:**W351-W355.

21. **Madej T, Lanczycki CJ, Zhang D, Thiessen PA, Geer RC, Marchler-Bauer A, Bryant SH.** 2014. MMDB and VAST+: tracking structural similarities between macromolecular complexes. Nucleic Acids Res **42:**D297-303.


A
B
C
D
E
F

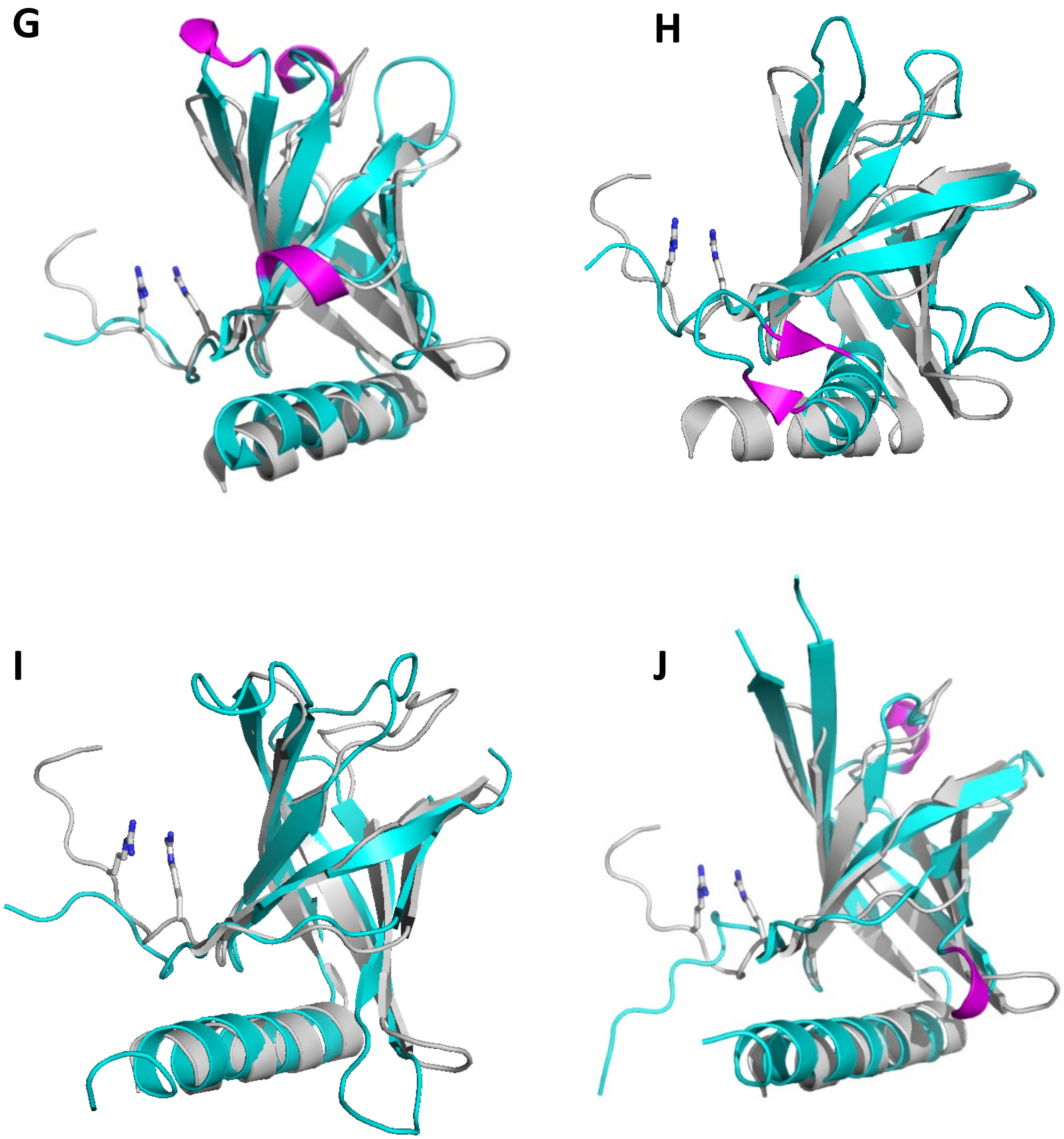
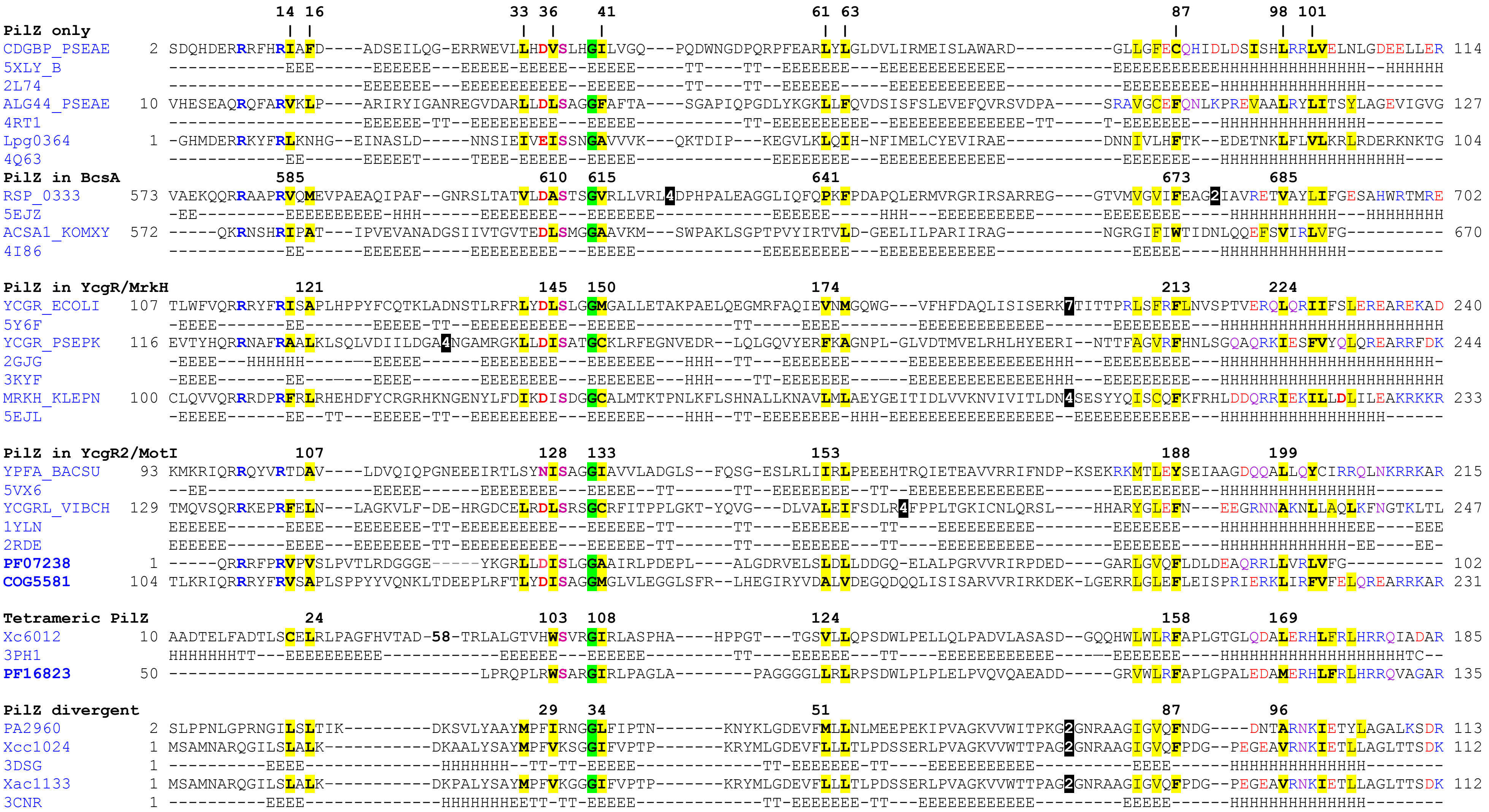


Figure S1. Structural superpositions of *P. aeruginosa* MapZ with canonical PilZ domains.

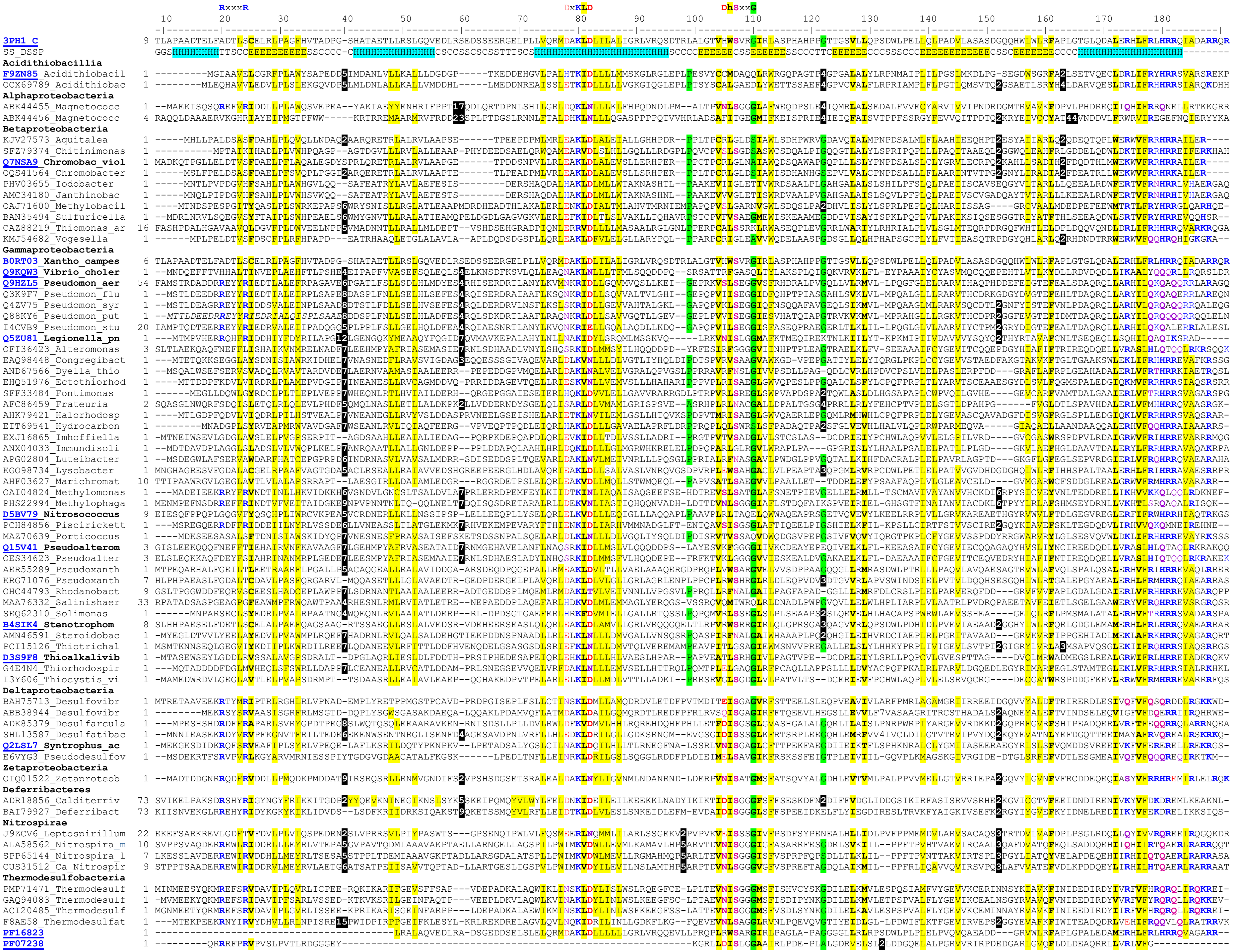
Stand-alone PilZ domain protein MapZ (PA4608) from *P. aeruginosa* (PDB entry 5XLY_B, residues 4-107) superposed against canonical PilZ domains listed in Table 1: 2L74 (A); 4Q63 (B); 4RT1 (C); 4I86 (D); 5EJZ (E); 5Y6G (F); 3KYF (G); 2RDE (H); 5VX6 (I), and 5KEC (J). In all panels, the MapZ structure is in light grey, aligned structural elements from other proteins are in cyan, structural elements that deviate from the MapZ structure are in magenta. The Arg residues of the MapZ RxxxR motif are in stick representation.



```
                                   14 16                       33 36    41                   61 63                                   87          98 101
PilZ only                          |  |                        |  |     |                    |  |                                    |           |  |
CDGBP_PSEAE      2 SDQHDERRRFHRIAFD----ADSEILQG-ERRWEVLLHDVSLHGILVGQ---PQDWNGDPQRPFEARLYLGLDVLIRMEISLAWARD----------GLLGFECQHIDLDSISHLRRLVELNLGDEELLER 114
5XLY_B             ------------EEE------EEEEEE---EEEEE-EEEEE--EEEEE-----TT----TT--EEEEEEE---EEEEEEEEEEEEE-----------EEEEEEEEEEEHHHHHHHHHHHHHHH--HHHHHH
2L74               --------------------EEEEEEE---EEEEEEEEEEE--EEEEE-----TT----TT--EEEEEE-----EEEEEEEEEEEE------------EEEEEEEEE-HHHHHHHHHHHHHHH--------
ALG44_PSEAE     10 VHESEAQRQFARVKLP----ARIRYIGANREGVDARLLDLSAGGFAFTA----SGAPIQPGDLYKGKLLFQVDSISFSLEVEFQVRSVDPA-----SRAVGCEFQNLKPREVAALRYLITSYLAGEVIGVG 127
4RT1               --------------------EEEEEE-TT--EEEEEEEEE---EEEEE-----------TT-EEEEEEEEEE--EEEEEEEEEEEEEE-TT-----T-EEEEEEE---HHHHHHHHHHHHHHH--------
Lpg0364          1 -GHMDERRKYFRLKNHG--EINASLD-----NNSIEIVEISSNGAVVVK---QKTDIP---KEGVLKLQIH-NFIMELCYEVIRAE----------DNNIVLHFTK--EDETNKLFLVLKRLRDERKNKTG 104
4Q63               ------------EE------EEEEET-----TEEE-EEEEE--EEEEE---------------EEEEEEEE---EEEEEEEEEEEEE-----------EEEEEEE---HHHHHHHHHHHHHHHHHHH----
PilZ in BcsA                   585                        610  615                        641                                   673       685
RSP_0333       573 VAEKQQRRAAPRVQMEVPAEAQIPAF--GNRSLTATVLDASTSGVRLLVRL4DPHPALEAGGLIQFQPKFPDAPQLERMVRGRIRSARREG----GTVMVGVIFEAG2IAVRETVAYLIFGESAHWRTMRE 702
5EJZ               -EE---------EEEEEEEEEE-HHH------EEEEEEEE---EEEEEEE-------------EEEEEE-----HHH---EEEEEEEEEE-------EEEEEEEE-----HHHHHHHHHHH---HHHHHHHH
ACSA1_KOMXY    572 -----QKRNSHRIPAT---IPVEVANADGSIIVTGVTEDLSMGGAAVKM---SWPAKLSGPTPVYIRTVLD-GEELILPARIIRAG----------NGRGIFIWTIDNLQQEFSVIRLVFG---------- 670
4I86               ------------EE------EEEEEE-----EEEEEEEEE---EEEEE-------------EEEEEEEEE---EEEEEEEEEEEE---------------EEEEE----HHHHHHHHHHHHH----------

PilZ in YcgR/MrkH                121                      145  150                        174                                   213       224
YCGR_ECOLI     107 TLWFVQRRRYFRISAPLHPPYFCQTKLADNSTLRFRLYDLSLGGMGALLETAKPAELQEGMRFAQIEVNMGQWG---VFHFDAQLISISERK7TITTPRLSFRFLNVSPTVERQLQRIIFSLEREAREKAD 240
5Y6F               -EEEE-------EE-------EEEEE-TT--EEEEEEEEEE--EEEEE----------TT-----EEEE--------EEEEEEEEEEEEE-------EEEEEEEE---HHHHHHHHHHHHHHHHHHHHHHH
YCGR_PSEPK     116 EVTYHQRRNAFRAALKLSQLVDIILDGA4NGAMRGKLLDISATGCKLRFEGNVEDR--LQLGQVYERFKAGNPL-GLVDTMVELRHLHYEERI--NTTFAGVRFHNLSGQAQRKIESFVYQLQREARRFDK 244
2GJG               -EEEE---HHHHHH-------EEEEE------EEEEEEEE---EEEEEEE---HHH--TT---EEEEEEE------EEEEEEEEEEEEEEEHHH---EEEEEEEEE---HHHHHHHHHHHHHHHHHHHHHHH-
3KYF               -EEEE-------EE-------EEEE-------EEEEEEEE---EEEEEEE---HHH----TT-EEEEEEE------EEEEEEEEEEEEEEEHHH---EEEEEEEEE---HHHHHHHHHHHHHHHHHHHHHHHH
MRKH_KLEPN     100 CLQVVQRRRDPRFRLRHEHDFYCRGRHKNGENYLFDIKDISDGGCALMTKTPNLKFLSHNALLKNAVLMLAEYGEITIDLVVKNVIVITLDN4SESYYQISCQFKFRHLDDQRRIEKILLDLILEAKRKKR 233
5EJL               -EEEEE------EE--TT---EEEEE-TT--EEEEEEEEE---EEEEEE----HHH--TT-EEEEEEEE-HHH-EEEEEEEEEEEEEEEEE--EEEEEEEEEEEE---HHHHHHHHHHHHHHHHH------

PilZ in YcgR2/MotI               107                      128  133                        153                                   188       199
YPFA_BACSU      93 KMKRIQRRQYVRTDAV----LDVQIQPGNEEEIRTLSYNISAGGIAVVLADGLS--FQSG-ESLRLIIRLPEEEHTRQIETEAVVRRIFNDP-KSEKRKMTLEYSEIAAGDQQALLQYCIRRQLNKRRKAR 215
5VX6               --EE-----------------EEEEE------EEEEEEEE---EEEEE--TT------TT--EEEEEEEE--TT--EEEEEEEEEEEEEE-------EEEEEEEEE--HHHHHHHHHHHHHHH--------
YCGRL_VIBCH    129 TMQVSQRRKEPRFELN---LAGKVLF-DE-HRGDCELRDLSRSGCRFITPPLGKT-YQVG---DLVALEIFSDLR4FPPLTGKICNLQRSL----HHARYGLEFN---EEGRNNAKNLLAQLKFNGTKLTL 247
1YLN               EEEEEE------EEEE---EEEEEEE-TT-EEEEEEEEEE---EEEEEE-TT------TT----EEEEEE---TT----EEEEEEEEEE--------EEEEEEE----HHHHHHHHHHHHHHEEE----EEE
2RDE               EEEEEE------EEEE---EEEEEEE-TT-EEEEEEEEEE---EEEEEE-TT------TT----EEEEEE---TT----EEEEEEEEEE--------EEEEEEE----HHHHHHHHHHHHH-EE----EE-
PF07238          1 -----QRRRFPRVPVSLPVTLRDGGGE-----YKGRLLDISLGGAAIRLPDEPL----ALGDRVELSLDLLDDGQ-ELALPGRVVRIRPDED----GARLGVQFLDLDEAQRRLLVRLVFG---------- 102
COG5581        104 TLKRIQRRRYFRVSAPLSPPYYVQNKLTDEEPLRFTLYDISAGGMGLVLEGGLSFR--LHEGIRYVDALVDEGQDQQLISISARVVRIRKDEK-LGERRLGLEFLEISPRIERKLIRFVFELQREARRKAR 231

Tetrameric PilZ                   24                      103  108                        124                                   158       169
Xc6012          10 AADTELFADTLSCELRLPAGFHVTAD-58-TRLALGTVHWSVRGIRLASPHA----HPPGT---TGSVLLQPSDWLPELLQLPADVLASASD--GQQHWLWLRFAPLGTGLQDALERHLFRLHRRQIADAR 185
3PH1               HHHHHHHTT---EEEEEEEEEE------------EEEEEE---EEEEEE---------TT----EEEEEE---TT----EEEEEEEEEEEE------EEEEEEE----HHHHHHHHHHHHHHHHHHHHTC--
PF16823         50 --------------------------------LPRQPLRWSARGIRLPAGLA--------PAGGGGLLRLRPSDWLPLPLELPVQVQAEADD-----GRVWLRFAPLGPALEDAMERHLFRLHRRQVAGAR 135

PilZ divergent                                           29   34                         51                                    87        96
PA2960           2 SLPPNLGPRNGILSLTIK---------DKSVLYAAYMPFIRNGGLFIPTN-------KNYKLGDEVFMLLNLMEEPEKIPVAGKVVWITPKG2GNRAAGIGVQFNDG----DNTARNKIETYLAGALKSDR 113
Xcc1024          1 MSAMNARQGILSLALK-----------DKAALYSAYMPFVKSGGIFVPTP-------KRYMLGDEVFLLLTLPDSSERLPVAGKVVWTTPAG2GNRAAGIGVQFPDG--PEGEAVRNKIETLLAGLTTSDK 112
3DSG             1 ----------EEEE--------------HHHHHHH--TT-TT-EEEEE-------------TT-EEEEEEE-TT----EEEEEEEEEEE----------EEEE------HHHHHHHHHHHHHH--------
Xac1133          1 MSAMNARQGILSLALK-----------DKPALYSAYMPFVKGGGIFVPTP-------KRYMLGDEVFLLLTLPDSSERLPVAGKVVWTTPAG2GNRAAGIGVQFPDG--PEGEAVRNKIETLLAGLTTSDK 112
3CNR             1 ----------EEEE--------------HHHHHHHEETT-TT-EEEEE-------------TT-EEEEEEE-TT---EEEEEEEEEEEE---------EEEEE------HHHHHHHHHHHHHH--------
```

Figure S2. Structure-based sequence alignment of canonical PilZ domains.

The protein names are linked to the respective entries in UniProt, the structure names are linked to the PDB. motif. Proteins are listed under their GenBank accessions, proteins from model organisms are shown under their UniProt identifiers and linked to the respective UniProt entries. Each sequence is followed by its secondary structure(s) derived from the PDB entries (H, α-helix; E, β-strand; T, turn). The sequence alignment was generated with VAST (21) and manually adjusted based on the alignments shown in Fig. S1. Conserved hydrophobic residues are shaded yellow, those forming the hydrophobic pocket are in bold; small (turn) residues are shaded green. In the RxxxR and [D/N]hSxxG motifs and in the C-terminal α-helix, positively charged residues (R, K, H) are in blue, acidic (D, E) are in red, N, Q, S and T are in purple. Also included are sequence and structure of the tetramer-forming PilZ (tPilZ) domain, the sequences of the eponymous PilZ protein (PA2960) from *P. aeruginosa* and its orthologs from *X. campestris* and *X. citri*, and the sequences of Pfam entries for PilZ (PF07238) and PilZ_2 (PF16823) domains.



```
                                        RxxxR                                                   DxKLD                      DhSxxG
                           10        20        30        40        50        60        70        80        90       100       110       120       130       140       150       160       170       180
                           -|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
3PH1_C                   9 TLAPAADTELFADTLSCELRLPAGFHVTADPG-SHATAETLLRSLGQVEDLRSEDSSEERGELPLLVQRMDAKLDLILALIGRLVRQSDTRLALGTVHWSVRGIRLASPHAHPPGTTGSVLLQPSDWLPELLQLPADVLASASDGQQHWLWLRFAPLGTGLQDALERHLFRLHRRQIADARRQR
SS_DSSP                    GGSHHHHHHHHTTSCCEEEEEEEEESSCCCC-CHHHHHHHHHHHHHHCSCCSSCSCSSTTTSCSHHHHHHHHHHHHHHHHHHHHHHHHSCCCCEEEEEECSSEEEEEESSCCCTTCEEEEEECCCSSSCCCEEEEEEEEEEECSSCEEEEEEECCCCHHHHHHHHHHHHHHHHHH--------
Acidithiobacillia
F9ZN85_Acidithiobacil    1 ---------MGIAAVELCGRFPLAWYSAPEDD5IMDANLVLLKALLLDGDGP-----TKEDDEHGVLPALHTKIDLLLMMSKGLRGLEPLPESVYCCMDAQQLRWRGQPAGTP4GPGALALYLRPNMAIPLILPGSLMKDLPG-SEGDWSGRFA2LSETVQECLDRLIFRYHRRSVARSREKP
OCX69789_Acidithiobac    1 --------MLEQHAVVLEDVLPLSLEKGQVDP5LMLDNLALLKALLVMDDHL-----LMEDDNREAISSLETKIDLLLLVGKGIQGLEPLPTSYSCALGAEDLYWETTSSAEF4GPVCVALFLRPRIAMPLFLPGTLQMSVTQ2GSAETLSRYH4LDARVQESLDRLIFRHHRRSIARQKDHH
Alphaproteobacteria
ABK44455_Magnetococc     1 --MAEKISQSQREFVRIDDLLPLAWQSVEPEA--YAKIAEYYENHRIFPPT17QDLQRTDPNLSHILGRLDQKLNLLLKLFHPQDNDLPM--ALTPVNLSGGGLAFWEQDPSLE4IQMRLALSEDALFVVECYARVIVVIPNDRDGMTRVAVKFDPVLPHDREQIIQHIFRRQNELLRTKKGRR
ABK44456_Magnetococc     4 RAQQLDAAAERVKGHRIAYEIPMGTPFWW-----KRTRREMAARMRVFRDD23SPLPTDGSLRNNLFTALDHKLNLLLQGASPPPPQTVVHRLADSAFITGEGMIFKEISPRIE4IEIQFAISVTPPFSSRGYFEVVQITPDTQ2KRYEIVCCYAT44VNDDVLFRWRVIREGEFNQIERYYKA
Betaproteobacteria
KJV27573_Aquitalea       1 -----MHLLPALDSASFDAHLPLQVQLLNDAQ2AARQRETRLALRVLAAPSE------TPEEPDPMLMRLEAKLDLALEIALLGHHPDR--PPLTPCRLGLDSIAWLSDHPWRVGDAVQIALMPNPDSALMLFLAAHIEEQHPT2ESYAIIARLQ2QDEQTQPLWERWVFRRHRRAILER----
SFZ79374_Chitinimonas    1 ----------MPTAIKIHADLPLVWHPQAGP--AGTDGVLLLRVLALLEAAP-PHYDEEDSAELQRWQAMEARVDLSLHLLGQLLLRDGPLPQVCPVSLSGDSASWCSDQALPIGQQGTLALYLSPRIPQPLLLPAQITAAEQL2GGWQLEAHFRLGDDELQDWLDKTIFRRHRREIFERKHAH
Q7NSA9_Chromobac_viol    1 MADKQTPGLLELDTVSFDAELPFLAQALEGDYSPRLQRETRLALRVLAAPGE------TPDDSNPVLLRLEAKLDLALEVSLLERHPER--PPCTPCRLGLNAIAWQDSQAWAPGQPLLLSLYPNPDSALSLCLYGRVLECRPQ2KAHLLSADIH2FDQDTHLMWEKWVFRRHRRAILER----
OQS41564_Chromobacter    1 -----MSLFPELDSASFDAELPFSVQPLPGGI2ARQERETRLALRVLAAPTE------TLPEADPMLVRLEAKLDLALEVSLLSRHPER--PPLTPCRLGLDSIAWISDHANHGSEPVLVALCPNPDSALLLFLAARINTVTPG2GNYLIRADIA2FDEATRLLWEKWVFRRHRRKAILER----
PHV03655_Iodobacter      1 -----MNTLPVPDGVHFSAHLPLAWHGVLQQ--SAFEATRYLAVLAEFESIS--------DERSHAQDALHAKLDLMLLWTAKNASHTL--PAAKEVIIGLETIVWRDSAALPLGAHGALALSLSHILPFSLQLPAEIISCAVSEQGYVLTARLLLQEEALRDWFERTVFRNHRRLVHAERGAQ
AMC34180_Janthinobac     1 -----MNQLPIPDGVHFSAHLPLVWHSVWQQ--SAFEATRYLAVLAEFESTS--------DERSHAQDALHAKLDLMLLWTAKNASHNL--PAAKEVVVGLETISWRDVAALPVGAHGALAISLSQVLPFPLQLPAEIVSCVGADQAYTVTARLLLKEEALRDWFERTVFRNHRRHIVHAERGAQ
OAJ71600_Methylobacil    1 ----MTNDSPESPGITYQASLPLSWRKEPAPS6WRYSNISLLRGLATLEAAPMDRDHEADTHLAKALERLENKLDIALTMLAHVTMRNIEMPAPQPVSLGARNVGWLSDQSLPA2DHVLISLYLSPRLPEPLQLPAHARIISSAG----GRCVAALMDEDPEFEEWMTRTLFRYHRRGLQARHQE-
BAN35494_Sulfuricella    1 --MDRLNRVLSQEGVSYFTAIPLSWHPEAELS6WMYGNVTLLRALATIEAMQPELDGDLGAGVGKVLERLEFKIDLTLSLVAKLLTQHAVRPSTCPVFVSAEGMEWISKEAAMEGDDIVISAYISPKLPQPLVLPAKIKSIQSESGGTRIYATFTHLSEEAQDWLSRTVFRYHRREVQQHSR--
CAZ88219_Thiomonas_ar   16 FASHPDALHGAVAAVQLDGVFPLDWVEELNPP5VMADNNTLLRALLMLDEPT--VSHDSEHGRADPIQNLERRVDLLLMASAALRGLGNLPPERPCALSSRKLRWASEQPLEVGRRLWARIYLRHRIALPLVLSGLMTEQRPDRGQFWHTLELDPLDQQVQNDLDRLIFRHHRRQVARKRQGA
KMJ54682_Vogesella       1 ------MPLPELDTVSFDSCFPLRFHPAPD---EATRHAAQLETGLALAVLA--APLDQDSDGSPLLQRLEAKLDFVLELGLLARYPQL--PPARPCRIGLEAVVWQDELAASPGDSGLLQLHPHAPSGCPLYLFVTIEASQTRPDGYQHLARLQ2RHDNDTRRWERWVFQQHRQHIGKGKA--
Gammaproteobacteria
B0RT03_Xantho_campes     6 TLAPAADTELFADTLSCELRLPAGFHVTADPG-SHATAETLLRSLGQVEDLRSEDSSEERGELPLLVQRMDAKLDLILALIGRLVRQSDTRLALGTVHWSVRGIRLASPHAHPPGTTGSVLLQPSDWLPELLQLPADVLASASDGQQHWLWLRFAPLGTGLQDALERHLFRLHRRQIADARRQR
Q9KQW3_Vibrio_choler     1 --MNDQEFFTVHHALTINVEPLAEHFTLPSHE4EIPAPFVVASEFSQLEQLS4ELKNSDFKSVLQLLEAQNAKLNLLLTFMLSQQDDPQ--SRSATTRFGASQLTYLAKSPLQIGQKVRVKLFL-EYPAAAIYCYASVMQCQQEPEHTLVTLKYDLLRDVDQDLLIKAALYQQQRLLRQRSLDR
Q9HZL5_Pseudomon_aer    54 FAMSTRDADDRREYYRIEDTLALEFRPAGAVE6PGATLFSLLSDLHLMDYES4RHISERDRTLANYLKVMNKRIDLLGQVMVQSLLKEI--GEPRKVSLSEGGVSFRHDRALPVGQLLVLRMVL-LPQGFGLELRARVIHAQPHDDEFEIGTEFEALSDAQRQLLARHILQKQAQQRRLARAGQ
Q3K9F7_Pseudomon_flu     1 --MSTLDEEDRREYYRIEDTLALEIRPLSAPE8DASPLFNLLSELHLSEFES4RQISERDRAIAAFLKSQNKRIDLLSQVVALTVLGHI--GEPQPVIISEGGIDFQHPTPIASGAHLSVKLVL-MPQALGLLLRARVTHCDRKGDGYDVGTEFEHLTDAQRQLLARYILQKQAQERRLAREQN
Q4ZV75_Pseudomon_syr     1 --MSTLDEAGRREYYRIDDSVALEINPLSAAD8DTSTLFDLLSELHVSEFES4RQLDERDRVLNSFLKSLSKRIDLLGEVVAHTALGKL--GAPQPVKLSEGGIQFNSQQAFAVGEQLSIKMVL-MPQAAGLMLRARVSQCDTL2GNFYISTEFVNLPDAQRQLLARHVLQRQAQQRRQALEQG
Q88KY6_Pseudomon_put     1 --MTTLDEEDRREYYRIEDRIALQISPLSAAE8DDSPLFNLLSELHLADFES4RQLSDKDRTLAAFLRAQNKRLDLLSAVVGQTLLGEV--GEPLPVVISEGGIESVHATQIAPGTRVKVKMVL-MPRAHGLLLRGKVTHCDPR2GGFEVGTEFIDMTDAQRQLLARYILQRQQQRRQQLELN
I4CVB9_Pseudomon_stu    20 IAMPTQDTEERREYYRIEDRVALEIIPADQGQ5PLPPLFSLLGELHQLDFEA4RQIAESNRTLANYLKVQNKRIELLGQALAQDLLKDQ--GAPQPVILSEGGISFASEQPRAEGERLTLKLIL-LPQGLGLVLAARVIYCTPM2GRYDIGTEFEALTDAQRQLLARHILQKQALERRLALESL
Q5ZU81_Legionella_pn     1 ---MTMPVHERRQHFRIDDHIYFDYRILAPG12LGENGQKYMEAAQYFQGID7QVMAVKEPALAHYLNLLNAKIDYLSRQMLMSSKVQ-----LRKVNISLGGMAFKTMEQIREKTNLKIILYT-KPKMIPIIVDAVVVYSYQQ2THYRTAVAFCNLTSEQEQLLSQHILQAQIKSRAD-----
OFI36423_Alteromonas     3 SLTLAEKQAQFNEFFLISHAIKVNMRELNADF7LEEHMPYAFRIASEMASIE7RNLSDHAADLVNYLSHQSRKIDLMMSYILHQQDDPD--YRFESIRFGGGVIINHPTPLEAGTRAELKLFV-SEEAAAIFCYGEVITCQQEPDGYHIAFIFTRIREQDQELLVRASLHLQTQQLRKRSQQK
EAQ98448_Congregibact    1 ---MTETQKKSEGGLAYSDNISIAWRKIDHEL7VNASNEDFLRAVSVIGDAG3EQQESSSGIVAQEVARLDLKVNLLLDLVGTLIYHQLDIPDTSPVRVSAAGVAWKGD-VPEPGATIYLELYIQRGLPKPLCCYGEVVSTAEDFAKGTAKVKFTGLTGAAKSWLEKLIFRHHRREVAFKRSSG
AND67566_Dyella_thio     1 --MSQALWSEFSERVSVADQLRVAVTARDVDE7LAERNVAAMASIAALEERR--PEPEDDGPVMQELARLDAKLNALVELVGRALQPVGSLPPRRAVRFNSLGIVVPSDLLPAG-QDLCVRLHPDVCPSLVLELPASLERPFDD---GRAFLAFRPLGEAHADALERLVFRTHRRKIAETRQSL
EHQ51976_Ectothiorhod    1 ----MTTDDPFKDVLVIRDRLPLAMEPVDGIP7INEANESLLRVCAGMDDVQ-PRRIDDAGEVTQELLRIESKVNLLVEMVSLLLHAHARIPPPVPLRISAEGLGWVQPESLPPGQALCLSFYLCPQFPRPLTLYARVTSCEAAESGYDLSVLFQGMSPALEDGIQKMVFRRHRRSIAQSRSRR
SFF33484_Fontimonas      1 ------MEGLLDQWLGYRDCLPLTLEPLVEPP7WHEQNLRTLHVIAILDERH--QRGEPGGAIESEIERLHQKLDVVLELLGAVVRARRGDLPTRPVRLSREGLSWPVAPDSPA2TQWLASLDLHGSAPAPLCWPVQILGVHE----GEVCARFVAMTDALGAAIERLVFTRHRRSVAGARSPG
AFC86459_Frateuria       2 SQASGLNWQRFSDQISLETQLRLQLEVLPHDF5QMQLNASLLETLLALDKPK2LLVDDERNDYSGELQLISARLDLVMAMLGRISPSVE--ASRHPLRLNACGALLDPALTGSG4PRRLLRLYFEHCPTVPLELSGTLDPAHPG------FVGLDTLSPAVHDALERLVFMHHRRQVAASRGK-
AHK79421_Halorhodosp     1 ----MTLGDPFQDVLVIQDRLPILHSTVEALP7VNEANEGLLRVYVSLDDASPRVNEELGSEISHELARIETKLNVILEMLGSLLHTQVKSPDPVTMRISAEGLGWVQAERLEPGQMLRMHWHLCPQFPRPLELYGEVASCQAVADGFDISVGFRGLSPLLEDGIQKLVFRRHRRSVAQSRAR-
EIT69541_Hydrocarbon     1 --------MNADGPLSYRVEAPMRWVAVDGAF7WSEANLRVLTQIAQFEERG--VPVEQPTPQDLEIQRLHAKLDLLLGAVAELAPRFLDRPPRQPLQLSWRSLSFPADAQTPA2SFGLVEVHLHALVLQPLRWPARMEQVA-----GIAQAELLAANDAAQQALERHVFQQHRRAIAAARRS-
EXJ16865_Imhoffiella     1 -MTNEIWSEVLGDGLAVSLELPVGPSERPIT--AGDSAAHLLEAIALIEDAG--PQRPKDEPQAPDLQRLEVKIDLLTDLVSSLLADRI--PRGTPVTVSADGLVLSTCSLAS--DCDRIEIYPCHWLAQPVVLELGPILVRD----GVCGASWRSPDPVLRDAIGRWVFRIHRREVARRRMQG
ANX04033_Immundisoli     1 --MDTDAVDPLAGGLSLADSLVLVWQPLKELP7ANRQAATLLALLGNLDEPP--PPLDDAPQLAAHLHRLDQKLDLLLGLMGRWHKRELELPDPQPARLGLRGVLWRAPLPA---GPGRISLYLHAALPSALELPATLLPLGD----GKVLAEFHGLDAAVTDAIERYLFRRHRRAVAQAKRPA
APG02804_Luteibacter     1 --MSDEGWLAFSERVAWDARFHATCEPGPRPD6INDRNASVLVAVSALMDRR-SDISEDDSPLTQEVARLDAKLNVLMEIVNRLLLPQSGLPPRIALRFNASGAVLPWDLGPLPPVGQTALLKIHFDACRALPLELPAVRLAGPTD---GKGFLGFEGLSEPVRDGIERLVFRQHRRQVAEARAHA
KGO98734_Lysobacter      1 MNGHAGRESVFGDALACGELRPAAFVAGTGDA5ACLRSEALLRAIAVVEDSHGREEPEERGLHDLAVQRIEAKLDLLSALVASLVNRQVGSDPVRPLEWSAHGACLVLPEAPTA3QPGMLRVRPCDWLPETLELPATVVGVDHDGDGHQLWLRFIHHSPALTAALERHLFRIHRRAVAESRRPR
AHF03627_Marichromat    10 TTIPAAWRGVLGEGLAVTLVLALAPSRRAPT--LAESGIRLLDAIAMLEDGR---RGGRDETPSLELQRLEAKVDLLMQLLSTWMQEQL--PAPVSATLSAEGVVLPAALLET--TDDRLEFYPSAAFAQPLVLGLEAVCELD----GVMGARWCFSDDGLREALGRWVFRMHRRAVARARR--
OAI04824_Methylomonas    1 ---MADEIEEKRRYFRVNDTINLLHKVIDKKH6VSNDVLGNCSLTSALDVLA7PRLERRDPEMFEYLKIIDTKINLIAQAISAQSEEFSE-HDTREVSLSATGLAFSNETPIEVGELLELRMLL-TSCMAVIVAYANVVHCKDI6RPYSICVEYVNLTEDDRELLIKHVVKKQLQQLRDKNEF-
PHS22994_Methylophaga    1 MENMPEFNSDRREFFRINDTVFVEITAIDGKE9NPVPNNTNLTQ-QQLNELT7DQISQSDRETARALRLLDEKVNLIASTIQHQQNVADH-LDTTDVNLSGGGIAFLSTDQFAIKSPVEIRIEL-GHSGTFINAIANVIACDKL6TPYYLRLAFSHMSEYDRNLLVKHTLSRQAQALRISQK--
D5BV79_Nitrosococcus     9 EIESQFPPQPLGQGVFYQSQHPLIWRCVKPEA5VCRDNERLLKLLNSSIPSP-LELLEEQPLLYSEIQRLELKVDLLLEQIGLLLAQQAPLPAAVPLRLTAQELQWWAQEAPRSGETVQVEVYLKEELRRPLVLLGRVKAREAETHGYRVWVLFTDLGEGVREGLERFIFRWHRRHIAQTRKGS
PCH84856_Piscirickett    1 ---MSREGQERRDFFRIDDEIILNYRLVSSDE6LLVNEASSLTLATGLEKMK7RHVEKEMPEVARYFTHIENKIDLIARHVMMNADGLFT-ENTQAVSISGSGLAFTIQEELPIGSHIELKFIL-KPSLLCIRTFSTVVSCIRE2GQYKIAVEFSKLTEGDQDVLIRHVVQKQMNEIREHNE--
MAZ70639_Porticoccus     1 ----MDKSEESASALSFTDNISIAWSKIDYQP7VNESNESFPRAVSAISEFSKETSDSSNPALSQELARLDLKLNLLLDLVGQLIYSQLDIPDISRVTVSSAQVDWQDGSVPEPGSIVFVQVYIQRGTPKPLCFYGEVVSSPDDYRRGWARVRYLGLSESVQVWLDKLIFRHHRREVAYRKSSS
Q15V41_Pseudoalterom     3 GISLEEKQQQFNEFFTIEHAIRVNFKAVAAGF7LGEHMPYAFRVASEATAID7RNMGEHAVELANFLNAQSRKIDLLMSLVLQQQDDPS--LAYESVKFGGGGIIVKCDEAYEPEQIVELKLFL-KEEASAIFCYGEVIECQQAGAQYHVSLIYNCIREEDQDLLVRASLHLQTAQLRKRAKQQ
OES34623_Pseudoalter     3 ELSLEQKKAQFDEYFSIAHRINANMRPLGEDE7LEESMPYAFRIASEMAAIE7RNLSDHAESLADYLNHQSRKIDLLMMSFVLHQQDEPE--FRFKTVKLGGGGVVIESKEALDVGAKAELKLFL-DAEAAAIFCFGEVITCEQVEDRYHIAFIFNTIREQDQELLVRASLHIQTQQLRKRAKEK
AER55289_Pseudoxanth     1 MTPEQARHALFGEILTLEETRAARFLPGALLP5ACAQGEALLRALAVIDDGA-ARSDEQDPQGEPALLRMEAKLDVLTLLVAELAAAQERGDPRQPLVWSARGVELVVSDPPAAGQQGLLRMRASDWLPTRLLLPAQVLAVQAESAGTRVWLAFVQLSPALQSALERHVFRIHRREVAQLRRER
KRG71076_Pseudoxanth     7 HLPHPAEASLFGDALTCDAVLPASFQRGARVL-MQQASETLLLGLAVAEDTR-GEDPDERGELPLAVQRLDAKLDLVLGLLGRLAGRLENPLPPCPLRWSHRGLRLDLEQPVDV3DTGVVRLAPVSWINDSIELPVTVLDQQHSESGQHLWLRTGALEPGYAEALERHLFRMHRRQIAEARRAS
OHC44793_Rhodanobact     9 GSLTPGGWDDFEQRVSCEESLHADCEPLAWPP7LSDRNANTLAAIAALEERR-ADTGEDDSPLMQEMLRMDAKLTVLVEIVNNLLVPGSVLPPRQLLRFNALGAILPAGFAPAD-GGLLLRMRFDLCRSLPLELPARVERQFDD---GRVFVVFAPLGDALGDAIERLVFRHHRRKVAGARQPP
MAA76332_Salinishaer    33 RPATADSASPGEAGPGFEAWMPFRWQAWTPAA4RHESNLRMLRVIATLETRE--NEPAEDDPLAQEFARLHDKVDLMLEMMAGLYERQGS-VSSRQVQMTWRQLRLDNADLPWPGVQVLLELWLHPLIARPLVLAATRLPVDRQPAEETAVFEIETLSGELGAAWERYVFRQHRREVSRGRQSR
SEQ62310_Solimonas       1 ------MNPARSECLSYEDRLPVALRPAATNG4WQEQNLRVLAAIATLDERP--RL-DPDSGTGAEFERLHRKFDVMIELLGALLRTQSSLPPQPQMVRLSSEGLSLPLSEAAPS2SLQEVELHLHACAPSPWRWLAEVSSHEA----GELQLRFLPMSMALATALERHVFTRHRRSVADARSPA
B4SIK4_Stenotrophom      8 SLHHPAESELFDETLSCELALPAEFQAGSAAG-RTSSAEGLLRSLALVEDSR-VDEHEERSEASLQLQRLEAKLDLAMVLLGRLVRQQGQELTLRPVRWSRRGIRLQLGPRSGA3QAGVVRLQPSDWLPDHIDLPVEVIAEAAD2GGHYLWLRFQRLGDGLEMAMERHLFRLHRRQVAEARRAR
AMN46591_Steroidobac     1 -MYEGLDTVVLYEELAYEDVLPVAWMPLRQEF7HADRNLRVLQALSALDEHGTIEKPDDNSPNAADLLRLELKLNLLLDMVGALLVNSQSRPQASPIRFNALGAIWHAAAPLPQ2QHGIELEIHVRDCIAEPLRLPGRITAVAAD---GRVKVRFIPPGEHIADLMEKLAFRKHRRQVAGARQRT
PCI15126_Thiotrichal     1 MSMTKNNSEQLGEGVIYKDIIPLKWRDILREE7LQDANEEVLRFITTLDDFHVENQDELGSASGSDLSRIEFKLNLLLDMVTQLVEREMAMPEAVPITLGSAGIEWMSNVVPREGALLELSLYLHHKYPRPLIVIGEVLSVTPI2GIGRYLVRLA3MSAPVQSGLEKIIFRQHRRSIAQSRRGS
D3S9F8_Thioalkalivib     1 -MTASEWSEYLGDDLRVSSALAVGPSDRALT--DPGLAQRLIESLDLFDDTH-PRSIPHEDESAPEIQRLEHKLDLLLHLVAESLHPAR--PAPVEATLSRHGIVLPAGTVPE--DCDRLEIYLSRLLPQPCVLGVESPTTAG----DVQLMRWADMEGSLREALGRWIFRLHRREIADKRQKV
G4E4N4_Thiorhodospir     1 ---MQTADDDDFDGLMVHEQLSFSWRLLDAPP7LCEANEALLRVCATLDDAM-PRLSNEGSEVVQELVRFDAKLNLLLEMVSELLHTTRQLPQMTPLELGAQGLRFPCAQLLAPPSLLLLDVYACPQFPKALRLFARVLDGQEDLEGYRIMARFEGLSTAMTEGLEKLVFRRHRRSIALRKHKL
I3Y606_Thiocystis_vi     1 -MAMEDWRDVLGEGLAVTLELPVAPSDRMPT--TSDAASRLLEAIAVLEAEP--QGHAKEDPVTPELARLELKIDLLMDLVTTLLADKI--PRRSRVGLSGDGLVLPATVLTS--DCERIEVFPCHWLAQPLVLELSRVQCRG----DECGATWRSPDDGFKEVLRRWVFRMHRREVARRRFQS
Deltaproteobacteria
BAH75713_Desulfovibr     1 MTRETAAVEEKRTYMRIPTRLRGHLRLVPNAD-EMPLYRETPPMGSTPCAVD-PRDPGISEPLFSLLCTINSKLDMLLAMQDRDVLETDFPVTMDITEISGAGVRFSTTEELSLEQPVEAVIVLARFPMRLAGAMGRIIRREEIDGQVVYALDFTRIRERDLESIVQFVFQSQRDDLRGKKWD-
ABB38944_Desulfovibr     1 --------MEKRSYSRVAASISGRFRLMSGTD-DQPLYSGWSGASAKDAEQA-LQQAKLPDAMVQFLATMDAKLDAILGQMQRDTLREDFPFRLRVSQISGAGIRFFTQEEVLHEGSLLEVVLF7VASAAGRITRCSTHADALS2AQNEYALEFTVINDSELEQVIGFVFDQERRIIRQHRWE-
ADK85379_Desulfarcula    1 ---MPESHSHDRDFFRAPARLSVRYGPDTPEG8SLWQTQSQLEEAARAVKEN-RNISDSLLPLLDVLRWLDFKVDMVLHHLRQREHDQHFPHLLETFDISGSGLGVASHGALALGQRLILAISLPNRPWRPIYARGEVVRDKKD2GQPRFGVRFSHIPEADQERLVRFTFEQQRRQLARRNQEA
SHL13587_Desulfatibac    1 --MNNIEASEKRDYVRVPFKGNVTFRILTEDE6EKENWSENTNRGLISENFD4AGESAVDPNLVRFLLHLEDKLDRILGLLGDKSRNGM--EVGSGIDISSSGLKFRTSRPLEEGQHLEMRFVV4IVCLDILGTVTRVIPVYDQ2KQYEVATNLYEFDGQTTEEIMAYAFRVQREALRSKRKQ-
Q2LSL7_Syntrophus_ac     1 -MEKGKSDIDKRQFSRVEAFIPLSYRLVPEQE-LAFLKSRILDQTYPKNPKV--LPETADSALYGSLCILNAKLDQILHLLTLRNEGFNA-LSSRLVNISGAGLCFTTPEKFAEGDIIEIKTFLSPHKNRALCLYGMIIASEERAEGYRLSLSFVQMDDSVREEIVKFVFERERELLREKRGS-
E6VYG3_Pseudodesulfov    1 --MSDEKRTFSRVPVRLKGYARVMRNIESSPIYTGDGVGDAACATALFKGSK------LPEDLTNFLLEINRKLDRILGSLSQGGLRDDFPLDIEIMELSAVGIKFRSKSSFEIGTPLEIVIIL-GQVPLKMAGSKGIVRGIDE-DTGLSRFEFVDTLESGMEAIVQFVFQQQREQIRKTKM--
Zetaproteobacteria
OIQ01522_Zetaproteob     1 --MADTDDGNRRQDFRVDDLLPMQDKPMDDAT9IRSRQSRLLRNMVGNDIFS2VPSHSDGSETSRALEALDAKLNYLIGVNMLNDANRND-LDERPVNISATGMSFATSQVYALGDHLEVTVMLPALPPVVMELLGTVRRIEPA2GQVYLGVNFVFRCDDEQEQIASYVFRRHREMIRLELRQK
Deferribacteres
ADR18856_Calditerriv    73 SVIKELPAKSDRSHYRIGYNGYFRIKKITGDP2YYQEVKNINEGIKNSLSYK5SKEIPQMQYVLWYLFELDNKIDEILEILKEEKKLNADYIKIKTIDISGGGFSFFSESKDFN2DIFFVDGLIDDGSIKIRFASISRVVSRHE2KGVICGTVFEEIDNDIRENIVKYVFDKDREMLKEAKNL-
BAI79927_Deferribact    73 KIISNVEKGLRREHYRIDYKGKYKIKLIDVDS--LSDFKRIIDRKSIQAKST9QKETSSMQYVLRFLLEIDTKLDLVLESLSNKEIDLEFM-EVDAIDISGGGLSFFSPEEIDKELFLYIEGDIRESLTRVKFYAIGKIVSEFK2RGYIYGVEFKYIDNELREDIIKYVFEKDRELIKKSIQS-
Nitrospirae
J9ZCV6_Leptospirillum   22 EKEFSARKREVLGDFTVFDVLPLVIQSPEDRN2SLVPRRSVLPIYPASWTS---GPSENQIPWLVLFQSMEERLNQMMLILARLSSGEKV2PVPVKVEISSSGIVFPSDFSYPENEALHLLIDLPVFPPMEMDVLARVSACAQS3TRTDVLVAFDPLPSGLRDQLLQYIVVRQREEIRQGQKDR
ALA58562_Nitrospira_m   10 SVPPSVAQDERREWLRIDDRLLLEYRLVTEPA5GVPAVTQDMIAAAVAKPTAELLARNGELLAGSPILPWIMKVDWLLEVMLKAMAVLHP5ARVTDVNISGGGIGFASARRFESGDRLSVKIIL--PPFTPVHTVAKVIRCAAL3QAFDVATQFEQLSADDQEHIIQHIIRTQAERLRARRQQT
SPP65144_Nitrospira_l    7 LKESSLAVDERREWIRVDDHLLMEYRLTSESA5STPPLTDEMIAAAVGKPTADLLARSGDALATSPILPWIMKVDWLMEVLLRGMAHMQP5ARLTEVNISGGGISFVSPRPFQPSDLLTLKLIL--PPFTPVNTTAKVIRTSPL3PGYLIATQYVELAPDEQEHIIRHIIQTQAERLRARRASA
CUS31512_Ca_Nitrospir    9 STPPTSAADERREWIRIDDRVLMEYRVLAETG6ATSATPEIISAVVTQPTAD-LLARTGESLIGSPVLPWIMKVDYILEVILNSLAMTHP5ARPTDVNLSGGGVGFVSPREFTAGDQLAIKMIL--PPFRLIQAVVQVIRSVPQ3LAFVVATEFVDLKPDDQEYLIRHILHTQAERLRARRAAT
Thermodesulfobacteria
PMP71471_Thermodesulf    1 MINMEESYQKMREFSRVDAVIPLQVRLICPEE-RQKIKARIFGEVSFFSAP---VDEPADKALAQWIKLINSKLDYLISLWSLRQEGFCE-LPLTEVNISGGGMSFISHVCYSKGDILELKMVLESPQSIAMFVYGEVVKCERINNSYKIAVKFINIDEDIRDYIVRFVFHRQRQLIRQKREI-
GAQ94083_Thermodesulf    1 -MVMEEKYQKMREFSRIDAVIPFSARLVSSED-KTKVKSRILGEINFAQTP---VEEPLDKVLAQWLKVINAKLDYLINLWSLKEEGFSC-LPTAEVNISGGGMSFISDIPYNKGDILELKMVLETPSPVALFLYGEVIKCEALNSGYRVAVQFVNIDEDIRDYVVRFVFYRQRQLLRQKKEV-
ACI20485_Thermodesulf    1 MGNMEETYQRMREFSRVDAVIPLGVRLISSEE-KPRIKARISGEINFARPP---LDEPADKALAEWIKMINSKLDYLINLWSFKEEGFSS-LATTEVNISGGGMSFVSDLSYNKGDILELKAVLESPSPVALFLYGEVVKCEMINNAYRVAVKFINIEEDIRDYIIRFVFHRQRQLLRQKREI-
F8AE58_Thermodesulfat    1 ---MTEKPEEKRNYIRVYDHVLLRLNPISRE15PWIDPIRPPGEIFKLESYL-KRLREKDRELAGVLQVLNQKIDRILINLLGDKFLKG--FQEVEVNLSAGGLRVNLPQEVQVGTIYEIDLGL-LPDWIFLKTFGEVIRVEPS2GGYEVAFKFIWITEADQDRLVEHIFRQQVLQLRATRRVK
PF16823                  1 -----------------------------------------LRALAQVEDLRA-DSGEDDSELSLELQRMDAKLDLLLLLGRLVRQQEVLLPRQPLRWSARGIRLPAGLA-PAGGGLLRLRPSDWLPLPLELPVQVQAEADD---GRVWLRFAPLGPALEDAMERHLFRLHRRQVAGARR--
PF07238                  1 ---------QRRRFPRVPVSLPVTLRDGGGEY------------------------------------------------------------KGRLLDISLGGAAIRLPDE-PLALGDRVELSL2LDDGQELALPGRVVRIRPDEDGARLGVQFLDLDEAQRRLLVRLVFG-----------
```

Figure S3. Sequence alignment of representative tetramer-forming PilZ domains (tPilZ). The top line shows the positions of c-di-GMP binding RxxxR and [D/N]hSxxG motifs and the tetramer-locking [DEN]x[KR]h[DEN] motif. The second and third lines show the sequence and secondary structure of the tPilZ protein from *X. campestris* (PDB: 3PH1). Proteins are listed under their GenBank accessions, proteins from model organisms are shown under their UniProt identifiers and linked to the respective UniProt entries. Conserved hydrophobic residues are shaded yellow, those forming the hydrophobic pocket are in bold; small (turn) residues are shaded green. Conserved positively charged residues (R, K, H) are in blue, acidic (D, E) are in red, N, Q, S and T are in purple. Two bottom lines show the sequences of Pfam entries for PilZ and PilZ_2 domains.

A

```
                                 --HHHHHHHHHHH----EEEEEHHHHHH-----EEEEEE -----EEEEEH-----EEEEE-----HHHHHH----EEEEEE----------------EEEEEEE--------EEEEEEEE------HHHHHHHHHHHH----
Borreliaceae
BB0733_B.burgdorferi 1 --MLLSRKIRDYGAKYRGKEIKMSTEINSFLNLRNTIEMRI-GSYTAFGVIYSISMDSLKLIFQEDTVLPALAK-NKNLGSIQLKKNSDSK-SSAAFFPFLSVKLLSASAYSSLNKEYNLLTLEFLSPAPEEIAIKVGKLLDLKLG 141
AEL69950_B.afzelii   1 --MLLSRKIRDYGAKYRGKEIKMSTEINSFLNLRNTIEMRM-GAYSVLGVIYSISMDSLKLIFQEDTVLPALAK-NKNLGSIQLKKNSDSK-NSAAFFPFLSVKLLSASAYSSQNKEYNLLTLEFLSPMPEEIAIKVGKLLDLKLG 141
ACH93662_B.duttonii  1 --MFFSKKIKDYEAKYRGKEIKMSTEINSFLNIKNSVEIKV-GTYVVVGVIYSISMNFIKLILQQDEILPILAQ-NGNLGNLNIKNYDNNV-VDNAFCI2LIVKLVSTSSYSIQNKEYNLLTLEFVSSMPEEILVKIGKLLDLKLG 142
ACH94956_B.recurren  1 --MFFSKKIKDYEAKYRGKEIKMSTEINSFLNIKNSVEIKV-GTYVVVGVIYSISMNFIKLILQQDEILPILAQ-NGNLGNLNIKNYDNNV-VDNAFCI2LIVKLVSTSSYSIQNKEYNLLTLEFVSSMPEEILVKIGKLLDLKLG 142
AAX18049_B.turicatae 1 --MFLSRKIKDYEAKYRGKEIKMSTEINSFLNIKNSVEIRV-GNYLAYGVIYSISMNAIKLILQEDKVLPVLAQ-NGNSGNIQFKSFDNVG--DDSFFI2LVVKLVNTSSCSVQEKEYNLLTLDFLSPMPEEVAVKIGKLLDLKLG 141
ACL34890_B.garinii   1 MLSSKKNIIKNYEVEYKGKEIKISTEIYNFLNIFSTVKMVI-DNYSMFGIIYSISMNYIRIIFKEKNETLLILAKNKNLCSIRLAENSYPK-NNKTYFHFLPVELLDIYAYSYQDKKYNLLTLKFLSSLPEEIAIKVGKLFELKFG 144

Spirochaetaceae
TDE0214_T.denticola  1 MAFAASQQLNRYYNLYKNIDVTFSKEVVSTLNFEPKQVFVRCSGGQWPCIINSASMTKAKIICGKKSGFLARLRSGITSVNIRFAFFDTEGKDSLSFF--VAAKLVGISSYEAGNQDLVLITFEYTQRAPDDLIEKLGILLEANIN 144
TP_0086_T.pallidum  13 VEFATREQLNRYYDLYKDVDVTFSKDVMQALCFNARQVCVRSAGGQCSCVMNSVSMVGAKVILSRKSSLLESIQVEGASVSIRFSFFESDARDAVSFF--VTARVLGVEDYAQ-STELVVLSVAYTQRIPDMLIERLGLLVEANIS 155
Alkalispirochaeta    1 MALTTSQQITRYYQEFEHTEVTFTQEIIKALFLNTKQIFIKCLGYQWPCIVYSSSMTGARVITTLQPSLKAALSKSKNMVSLRFSFIQREKTDPLAFF--VTARITGFSPYGDPSKGLSFITLHYTQRPPDDLIERLGNLLSATIS 144
Marispirochaeta      1 MALITSQQLSRFFERYNDVNLTFTKDVVRATRLLQRNTHIKCLGDHWPCVLYSCSMKGAKVIASLSKSFYEKLKSANNLVTLHLAFAQEDKSTPLSFF--VTSKINGFTPYDKSKPNLNFLSLEYTQRPPDDLIVILGTLAEANIN 144
Oceanispirochaeta    1 MAIVTNQTIKELFNRFRSIEVTFNKEVSKATGLQPRKIFLKFKESARPCILYASSMEAARVIVSLPTAILQNLKENESNVALKLSFIKDEQ3IELNFF--IQGRLSSFTRYSDDQPDVYFCSINFSTQPPDDLIETLGFLLEANVN 146
Olavius_symbiont     1 MAFVTSQQLDAYYDKYKEIDVTFTKDVAKSLRMIQNQTHLKYKGGQRSCVIYSSSLCGSKIIIALPNKLLAKL-QTAKLVYLRYCFQKEDSNDIISFI--VHSRVIGFTQYGDN-KEIYFAQLEYTHRPPDALIEIIGAMLDANVN 144
Salinispira_pacif    1 MSVVTSQQISKLYDRYSEQDVLFNKDINRALLLNTKGIHLRCLGYQWPCILYSSSMTSAKLVVNVKTPIKKVIEQANRMVSLNYSFLQKDKANPISFF--MKAKVVSYAPYHPEKPDLNFLNLEFAHRPPDELIFRLGELLEAQTN 144
Sediminispirochaeta  1 MALMTSQQINRFYQHYREIDVTFNKDVIRATGLLAKHIFLKFLGSQIPCIIFSASMTGAKVIANISSETFERLRQANNLVSLRLSFQESDKNDPLAFF--VASKISGFNPYNKEHPTLNFVNLEFTQRPPDDLIGILGTLLEASVN 144
Spirochaeta_afric    1 MSITTSTQITRYYEDFKDIDITFTKEVIRASLLNTKQIYLKLLGYQWPCIIYSSSLSGAKIIANITPQLKKAMQQSKSGISLRFSFLQRDKRDPISFF--VPCKVAGFTPYQSAKGNLHFINMNFTQRPPDDLIEILGRILEANSA 144
Spirochaeta_therm    1 MAIVTSQQITRYYNQFKHIEVTFTKGINAAIGLVPTQTYVKCLDEQWPCIIYSSSMEGAKIILNNKLNFMDKLKEAKNLVSLRFCFQEREKHQPLSFF--VSAKVAGITPYGKQDASVSFLTLSYTHRPPDDLIEILGTLLEANIN 144
Treponema_primitia   1 MSVLTSQKIASYYERYKEIDVTFTKEIIQVTGLITQQVLLKCVGDFWPCVIYSSSFQGAKVVANIKSGLTEKLQQANNSVSLRLCFKSADSASPVTFF--VTARVAGYSPYGG-SSDTALFTLLFTQRPPDDLIEIMGRILDANIN 143
T.azotonutricium     1 MGVVTSQKITVYYDRFKTIDVTFTKEIIQVTSLITNQVFLKCGSDFWPCVIYSSSFQGAKIVANIKSGILHKLEQANQMASLRFSFKNPDTGSPVTFF--INTKSLGYSAYGG-SADVALFNLQFTQRPPDDLIEIMGRILDANVN 144
T.brennaborense      1 MGVVTSQQLLRYYELYRDKEVTFSKEVIKTLNLDPRQVYIKCTDSQWPCIINSTSFQGAKIIIGSKGGAYARLSQEKGTVNLRFCFLQADK-QPISFF--IGSRVVEITPYMN-SPDLAVVSLAYIQRPPDALIECIGKLLEANIN 142
T.caldarium          1 MGVLTSQQITKYYERYRTISVTFTKEIITATGLVTQQVYLKCLGEAWPCVIYSTSFEEAKVIINTKSGLYGKLQQANNLVNLRYSFKISDKTDPMFFF--VSCRVVGSAPYQG-SQDISMMTLQFTQRPPDDLIEIMGRLLDANIN 143
T.maltophilum        1 MGVMTSQQISRFYDSYRNTEIMFSKETIKALNLDPRQVYIKCNSGQWPCIINSSSLSIARIIIGTKGGAYAAIQKVNTPVSLRFFFSDPNG-QTMSFF--VNAHVSAMDPYMN-SKELAIATLAFNSKPPDDLIEIVGRLLETKDN 142
```

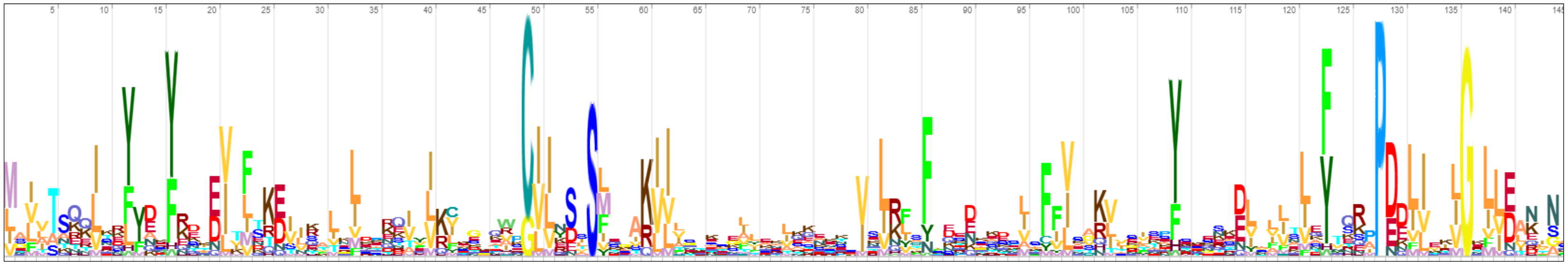


Figure S4. Sequence alignments of the newly identified xPilZ domains.

A. Sequence alignment of select representatives and an HMM logo of the PilZN3 domain, generated by the HMMer search at the EBI web site https://www.ebi.ac.uk/Tools/hmmer/. The top line shows secondary structure prediction by JPred. The sequences from *Borreliella burgdorferi* BB0733, *Treponema denticola* TDE0214, and *Treponema pallidum* TP_0086 are linked to their UniProt entries, sequences from other organisms are linked to the respective entries in NCBI protein database.

B

```
                            HHHHHHHHHHHHHHHHHHHH----HHHHHHHHHHHHHH----HHHHHHHHHHHHH------HHHH--HHHHHHHHHHHHHHHH-----------HHHHHHHHHHHHHH--
Spirochetes
TDE1318_T.denticola       1 MYILVVFFIVFLLFTAFSFTLPGRRFLDLIRFFTEGKDRGFLFSNLLLLWRTANYVGLEDKTRLFWSVAALDECIRFIARQVENKLDADVSQKMQILLNKLYDYRTKIEL
AEF80420_T.azotonutr      2 LPVLIIIIVVGGVVALVLARSKKGKSASWIQFYAKGKDSGFSFKEIELLRRLAVKANLEDPASLYWSQNQLDVCIRSLVRNMHLTGEDSD-TGSHDFLSKLYDYRKKIEM
AEJ19732_T.caldarium      2 LYFMLIIIATISLIILLTFRSDSKEKYSWLEFFAKGKDAGFSFGELQLLRKLAISAELEDPTTLFWSEKQLDRCIRELIKKAEVTGASRN-QETQDFLSKLYEYRKKIEF
EPF30613_T.maltophil      1 MYTAVIIVACILILIVLLRVI--FLISPCISFFAKGFNAGFDFKELRFLWPIADAGNPETPAALFESVALLDKSIAAVLKKAEKDKTTNT-QNVQSFLKRLYDCRTKVEL
AEF85949_T.primitia       2 VYILMALVPALAIVVFILHRPAGSSPNSRMLFYAKGKAAGFSRKEIGVLRRIAVRDEMGDQTAVFGDPELLGRCIRTLVQTMRAVGDRD--PAGQDFLARLYEYRKKIEL
POR05181_Alkalispiro      2 ELLISALALGLATVLLYLAFRTSGPGASWLRFYVRGKESGFSLAEVHLLRRVARERALSNPEALFWSVRTLETCLQGVLQQYHAAAPSRE-EERRRFLNKLFDFRAQLEF
ORC34506_Marispiroch      1 MIVMLVVFLVIVLLFILFLRHAGGGSFPWIHFYTKGKESGFSFREINLLRKVAVHNHLDNPISLFWSIRQLDRSIRGVMLRFRTEGREND-ESSILFINKLFDFRKKVEL
RDG32267_Oceanispiro      4 LYTVLTIFIIVSSIPLLLAVRKGLAKVDWVEFYSMGKDAGFSFKEIGLLRKVAVVNKHQKPSSLFWSIDVLDKSLAGIHKQIETESDPDEKDYLEYLLKKLYSYRKNVEF
VDB00783_Olavius_end      3 IIILLILTAILFVVIRMPKQTSATGIRGRLDFYLKGMDAGFSLSEIKVLWVGIKNSNFEKPSRIFGSIEALNALIGELIVNKKFLGRDGL-ESETQALKNLFAYRKDIEL
ADK80632_Sediminispi      2 NWTILAIVCGLIFLLIALMRRAGGGSFPWIHFYVKGKEAGFSFREVNLLRRIAVEARLKNPTSLFWSIGQLDRSIRGVVTKFRSQNLMDD-FNASQFLLKLFNFRRQVEL
ADN01812_Spirochaeta      1 MTVFFIIMGISLFVFLIGFQRMGGFTFPWIKFYVEGKEAGFSFSDLNLLRRLAVDNHLKDPTALFWSEKTLNLCIRNTIIKFRNQGREHD-PHSIEYIGRLFDYRKRVEF
Aquificae
Aq_1917_Aquifex_aeol     28 LLIIVLGALITVVSLPFFYSKLKEKQRIRENFFRRARDFDLTEEEASVLWKYVKET-VLNPNLVFENKAVFEKVVDRIVQNGN-----------PEEIKLISSIRMKLRF
ADY73882_Desulfuroba     23 VFVLIIVLIVLFFLLAGIIKEYLREKNLRISFFKEALERSLTKEEAAVLWNFSKKM-GRDPFLSLEFKAPFEKVVDLYIKTDP-----------NAKEEIIQDMRTKLGF
EDP76156_Hydrogenivi     28 LFFVVLAATLTVITIPYLWGKLSHAREVEKAFFMRGRSLGLTEEEIRLLWNYAKKF-PYDPQMVYENKPLFERVVSRIIHENI------------AGVRLIPSIRAKLRF
SNZ09612_Persephonel     21 IFAVVIVLLIFALIGFFLVF6FVSKRYLKKLFFRNGESYGLTKKELEILWKYSQKL-HKDPFLVIEYKAPFEKVIQAYIEENK-----------NYDEKMIRNMRKKLGF
AHE95351_Thermocrinus    22 LADVLLVLFGFVASVFVLIV7YLASRSVRREFEVVGKQMGLNEEEIALLYKCAST--LEEPNKVFYSKYVFERCVGKLVKES------------SDNIPIIVSVRKKLKF
Fibrobacteres
SHG10303_Fibrobacter     28 VWAALIISLVFVLIALLQAR6QKEVMFGTKTFDDKVAERGFSYNEKEILDKIIRKSPFENKDAILNSSGLFEQAVSTYYDVEGINEL------HETDLAVVESLRLKMNF
ACX73865_Fibrobacter     24 SLQILVIVLIFLAVVLIEVQ6ETDVMFGTDAFNEKIKAFDFTPKEVETLGDLVRSSKFENMDAVLNSSGLFEAAVSEYYRIRNVFSV------RDETLEAIAGLRRKMDF
KMQ51624_Chitinispir     63 IFLVMASGIGFLSFVYVRTR4KSAREASAYLFEQLIHNADLTRLEQQRLRTLLLYENIKDPQIIFQSISLYERCLDAYVQNFLSVSPS--SESVEHEEQVLGNLRKKMGY
Epsilonproteobacteria
SMC09305_Nitratirupt     27 IVVLFFLAIVLFFIIGTILQKRLKEKNIQHYFDIYAKERELNEREKETLWKYAHKME-RDPMLVLEFKAPFEKVMDLYIKTDP-----------HPDEELVRSLRKKLDF

Spirochetes                   --------EEEE-------EEEEEE----EEEEEEEEEE---EEEEEE---------------EEEEEEEE---EEEEEEEEEEE-------EEEEEE--
TDE1318_T.denticola       111 EEVQKKRRIESTHEIYIGQICIIFVPRETTVYGKLMANTKNELVFALFDASAERAEKVNWQNKAVRVYFWKQNDAGYVFSSEAVKAKKIEDRIEIYIKHS  210
AEF80420_T.azotonutr      111 DKPSIKNGISNSRQISDGQNLRVLVAGSGVFKSQIVKNTNQ-YITISRPTSTKLPGSFSWQNMKISAYFWREDDAGYVFDSEVLDEVFSKGIASLKISHS  209
AEJ19732_T.caldarium      111 EHPRYKKGITSSRSIGETQPIKILIDGMGVYKSQVVANSDR-FLTIQKPNIPPQKGYIVWKGKRLAIYFWRKDDAGYVFDSYVLDEAMARGMIVLQIAHS  209
EPF30613_T.maltophil      108 -DPNAKNGIQDTHYISKGQELRILVRGVGVFFARVMQNGRQLIISLPLENGELTVSGDKWVHKKIVVYFHRQDDAFYVFDTTVLRNLLFNGRNALVLAHS  206
AEF85949_T.primitia       110 -GAGQGQGITSSRQMSEGQNLRVVIGNSAMYYAKVIRSGWD-YLLITRPANPKLTTAPSWLGQKLAIYFWREEDAGYVFDCEVLDEVYFKGIAAIQVSHS  207
POR05181_Alkalispiro      111 RQPKYRLGISSTRRIEPGQPLKITVAGRGVIAARVSETARRYLAVRLSGEGGAVPAGFSWKNQRLGIYFWRRGDAGYYFETTVVAVPPNQKVPVLHLAHE  210
ORC34506_Marispiroch      110 SLPKYTLGLKTTRKLTTHQRLKITLPGVGTYNSSVVENLRR-YMAISYPEGPKLPPGASWRGQKINVYFWRAEDAGYVFETKVLEDFIEQKYPILHVAHS  208
RDG32267_Oceanispiro      114 KKPRYNRGIESTADISINQIVKLKVDVLGIFECSVLENNRN-YILVTYPKGPPVPVGFSWRDRTLNVYFWRREDAGYFFQSRIMEGYPDREFKTLRMSHS  212
VDB00783_Olavius_end      112 NRMKHYMGLKSTRSISIDQSLTIRV-SGVGVYSTTVVENKDNYLTITIPVGNPLPMGFSWRRSKLNVYLWRKDDAGYFFQTRILEKFYDKSNLLFHLKHS  210
ADK80632_Sediminispi      111 NLPKYTLGLKSSRKIMQRQRIKISLPGAGTFDSQVVENLRR-YLAVSYPEGPKLAPGFLWKGQLIQVYFWRVDDAGYTFETRVIEDYSEKQYPILHISHS  209
ADN01812_Spirochaeta      110 SLPKYRLGLRSTREIEAGQILRIILP3GRIFTSSVVENMRK-YIAISYPQSPH2PPDFSWMGREIHVYFWRKDDAGYFFTSRVLGDFKERKFEILHIDHA  211
Aquificae
Aq_1917_Aquifex_aeol      126 SSLPWFIPLTSTRDIEVYQTGVLVV-RNRRVDAYVYDKDEEFLYIALLEPVVVK------PGERVSFFFLRENDARYSFDATVEKVFNEGGRTVIVVRHT  218
ADY73882_Desulfuroba      121 DFVPYFVPLTSTKDIELFQPGKLYV-DNLSFDVALFDKDERFMYWAIIEEIPSNL---SLIEKKVKISFIRKNDGIYALEGTVKETYIDNGKLVLKIPHT  216
EDP76156_Hydrogenivi      125 DTIPWFIPITSTRDIDLYQTGKLVI-DGVYVDAAVWDKTETELHVVVLEALPRPV----RIGESVKFHFIRENEGRYSFETIIKDKYTEGDRLVLVLEHT  219
SNZ09612_Persephonel      124 DKIPPFMPLISTKDIDLFQTGTLIF-ENRVFPVSLYDKDEEYMYWYLIDQMPPFP---FKKGDTVRIRFTREDDAVYIVEGKIEDLFSEDGKYIIKIPHT  219
AHE95351_Thermocrinus     124 EHLPWFLPLATTRDIEVYQTGFVSY-KGKSYGAAVWEKTEEDLKIAILDRAPEFP----QPGEKVRFSFLREDDGRYYFEAEVLNTYLEGGKVILVVPHT  218
Epsilonproteobacteria
SMC09305_Nitratirupt      128 EVVSPYIPLITTKDIEVFQNGRMILSNNKALDVALYDKDEKYMYWLVIDGDLPLSTH---PGEYVKIVFIRQDDGIYSFEQPIAEILKESGKTIIKIPHT  221
Fibrobacteres
SHG10303_Fibrobacter      137 TASNPLAEIYSTRQFNVGDRIDLILENGTLIKRSDIVWRNEKEWAISYDGSDGPAS--SFVNRTIRIRWTRPDDAMYSTAISVKNADET---NNLVLPHS  230
ACX73865_Fibrobacter      133 TGSNPLSMVCSTRQFNIGDKVDLEFENGQSFKRIPIYERTEKVWSVKIDGSPSLAK--ALVGSHALVRWTRMNDAVYSIHLNVLAATP----DKVVFTHS  226
KMQ51624_Chitinispir      174 NFLAYEHPMVSTRNLSIGQKVSIVS2NKSVLIQNATVVKNRELSFVLQYNAEKEEVVNLLPGTQVNIVFTRYNDGAYGVNVKVKKLES----GTIELMHT  270
```

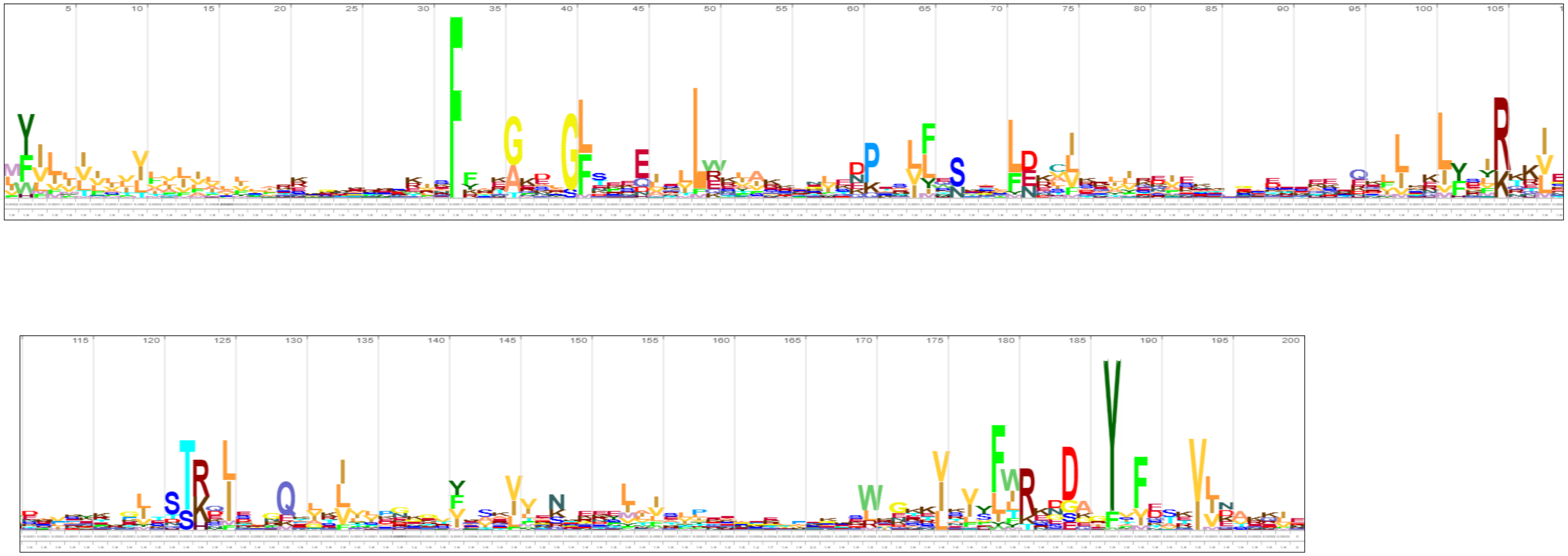


Figure S4. Sequence alignments of the newly identified xPilZ domains.

B. Sequence alignment of select representatives and an HMM logo of the PilZN4 domain, generated by the HMMer search at the EBI web site https://www.ebi.ac.uk/Tools/hmmer/. The top line shows secondary structure prediction by JPred. The sequences from *Treponema denticola* and *Aquifex aeolicus* are linked to their UniProt entries, sequences from other organisms are linked to the respective entries in NCBI protein database.

C

```
                                       -------------EEEEEEE------EEEEEE-----EEEEEEE------EEEE----EEEEEE------EEEEEEEEEE-----EEEEEEE------
Gobs_0077_Geodermatoph      1   MNTPGVDHPHRDGEVEVVLLGHRVSVSARVEAAD-AATVVVRPSVGEFAGQQVVRIGEEVQLVWHDAPDVRMVTAEVATVQHGAVPRWHLGVTAPAQP   97
Gobs_0078_Geodermatoph      1   MSIAGVDWPEPGGAAEVIPLARGVALPSSVEVSD-GTALLVRLVLADYAEKLKVQPGHRVEVLWGGPTGGRSLPAEVTEIRSDPELRWILQITGPAEA   97
Acel_0863_Acidothermus      1   MSRAGADTLSLNALVTLRIGDDEVDYPSRIEDIA4EDSVFVAAPT-GGSASLVASGVRTVELSWVSPRGRYQQPCEVVEFISGQPRLWRLRPTGPAAL  100
Acel_2035_Acidothermus      1   --MLVPPLVTGGPVTVRLVHRDGREFATRVESVS-PRVVVVAAPP-GANAALIASGTREIDLSWLSPRGRYEQRCELEHSA-GTSRQWRLRPLRPAVL   93
ACPL_2299_Actinoplanes      1   ------MELPEIGSEMFLALGEGVNFRSRLESVD-GEVFSVAAPL-ETAGPSGFLPGYEFDVFWVPPRSRVVMPVLLKSVTDGAPFRWNLFPTAEPVI   90
ACPL_4953_Actinoplanes      3   AAVSSTVELPAVGTPMFLVLGEGMNFRSRLEAVD-GDRFTVAAPL-ETAGPVVISAGQQFDVFWAPPSTRIVLPARLVAVSDEAPFRWSLTAAGLPQH   98
Krad_1675_Kineococcus      16   QWPGLNDRVWVELQMPDTLGGELIRLPTRVEDTG-EASLVVAAPG-FRGDLHVVAPGLPVTVAWAGTRGRSKQDFLIAEVVRRRVAAWDLAPCGEVVV  112
CCG01087_Blastococcus       2   SSVPGVDHPVEQTEAEVRLAGTNIAVTARVEVVH-EGVISVRPSAGDFAQDTVVKVGDPVEVFWRTDDGQRALPADVLEVQQGAVVRWRLAVTGVAEH   98
QBI19714_Egibacter          9   TMTANQIYPEINERVELARDEWDGPARSRVEDTA5TMALVVSAPRRAGSGAARAGPGDVVEITWANERGLCCADAELIAAEEEPVPTWALALPRGPEI  112
SCX58702_Klenkia_marina     1   -MTDGEDLPGPRSVLDVHVPDRDDLLISFVPEEQ2DGVLVVTVAQDRTGRRIRVPVGDTVELIWSSRRGPRARSARVVEVVTTGEPVWRLEPLGTVRD   98
SCX39608_Klenkia_marina     1   ------------MAVELPVTGALVDVVAGAEVLT-GHVALVRPGAVVVELPVEPVRTEPLRIYWAAPEGVHTVLARLGLRWP2TGVRWQLLPAASPER   86
RNM16143_Marmoricola         1   -------MESPEINDPVVLRDDELEYRSRVEDLA-PGLLTVARPL-GLPAEHRYEPGVALQVSWSSPRGIAVLPTRLVETRI2GVALWLLEVAGDGWV   90
RNM12486_Marmoricola         1   -----MEQHPDLGDPVVLRDPEGREHPSTVRDLG-NGLLVVAQPR-DVPADEVIGTGADVTVTWADADGV2VLPTRILAVHS2PVARWSLAVTGPAYT   93
CCH90514_Modestobacter       1   --MPGTDHPEERTVVDVLVGGSDSPLLSWVERVE-DGDIVVTVGQDRSQRRVRVADGDPLELVWRAASELRSLPARLVATEA3GEVCWRVRPTGPAAR   97
RKS75603_Motilibacter        1   ----MSALPELNTLVRVALGEGAPVLPSRVEGVE-GDDLLLAAPS-YVGDVVGPRVGSTVSVHWTSARGVCSVPVEFVAAERSGIKLWRVRLNGGVEF   92
```

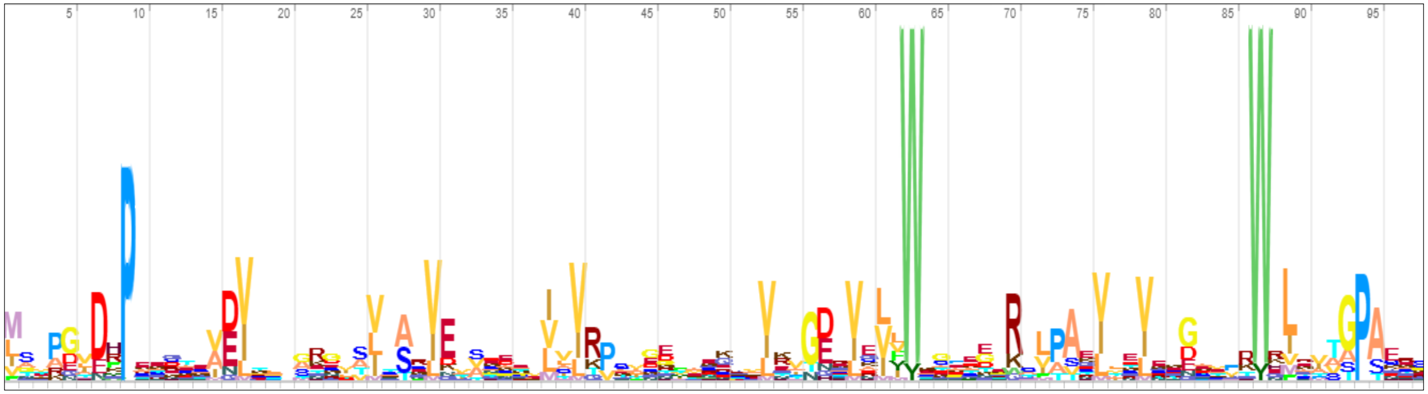


Figure S4. Sequence alignments of the newly identified xPilZ domains.

C. Sequence alignment of select representatives and an HMM logo of the PilZN5 domain, generated by the HMMer search at the EBI web site https://www.ebi.ac.uk/Tools/hmmer/. The top line shows secondary structure prediction by JPred. The sequences from *Geodermatophilus obscurus*, *Acidothermus cellulolyticus*, *Actinoplanes* sp. SE50/110, and *Kineococcus radiotolerans* are linked to their UniProt entries, sequences from other organisms are linked to the respective entries in NCBI protein database.

```
                                    HHHHHHHHHHHH---EEEE--------------EEEEE----EEEEE------------EEEEEEEEE---------EEEE-----EEEEE-----------EEEEE----------
Deltaproteobacteria
Deba_0512_Desulfarcula     25   QEELLAEALRPGRQVCLAVAADLVSDRIDVRPSMVHDIVKGGLLILAQTNPPLAKRAEGQTVEITFLIPHPDR-GAEYLRLGYKTKILRVIDGWKVDERLTDNIIVAPKPHRLERTT  140
Deba_3043_Desulfarcula     64   QEEMRRRAIRLRQLVDIIVDLDWSTDTLDVRASVIQEIDAKGRLILSQTSSPLLRSQTGKTVELTFLSQYPSQMGSRWLRVGYHTKLLGIVENYQLGPELADTVLVFDGPKKLVLST  180
ACN16127_Desulfobacter     11   IDLDIERILSPCLGVEIVFDLNSLT--PLVRSSIIYDVD2KQIFITAQPRILIHPNFVAKTIHLTTLVTVQG----SKKRIGIKCRPIQFDNNYRLSNDNRVGAIIFKYDPPITETT  122
KMY66548_Desulfocarbo      14   QSVLMSTCLRPGRPVELILDTSLDQNQIDVRPALLHDLDPKGRLVLSQTAPPLGRRWIGRQVEITFLARYLDVPGGRWLRVGYRAPIADLVADYRLDSTRVESVIIAPKPQKLQQTT  125
ROQ92034_Desulfosoma        1   -MTSPAKSLRPGAAVDIVLRVDFYREVMDVRRAMIYDVE-GTTLVVSRPTPDVKPDHVGRRCTVTFVVRENE----FKNRYGVPAKIVELKDNYTIRKGASVQAMILKVLGSVEPFN  111
EPR43376_Desulfovibrio      8   LAPFRERAAKSGR-LLLIVEEDLLRETIDAREVAVVDMD-AQRLELSQPEPPLPSSAAGRVMEIAAPPLPDEG--GAARPLGYHARLLEVN----------AESLVVSPPSGVRETS  110
SNS20366_Desulfovibrio      8   LALDPAVLLRAGAEAFIVLHKDALRERIDVRQTRVLDWE-GDVLFLAQTDPPVPREASGAQAEVVALAPACGE---ELRPMGYVARVLDVRPDYP-SAGLAVEAIAV3TPCDLFETS  121
OPY83705_Smithella          1   --MIKVPQLKVGTRLNIVFENEIMQSNAHFMKALVYDIQ-EGKIIISQTSPALNRNFLNRRIMVSFLAQMDG----RNLRFGFPATLIDLINDYKIVSEKNVEALVLQRDKAAEPLE  110
ABC78461_Syntrophus         1   --------MRVGLHLDIVLEEDLERDRVVARKAVVYDVV-ERLITISQTSPSFVRRHLGDRINASYVTTVGEGENKRPVRMGFPATILEYREDYPLASSNLVPAFVLQQDGKPVEIN  108
Omnitrophica
OGX18185_Omnitrophica       1   ------MRLNIAPGEEIHLLLDGNK----IRQSKIYQLLSNDFFSIEQTQPIIEKDYLNKILLLTYLTPKE-----RPGRFGFEARIQAVTSD---------ARIILHKLNDPAPCD   93
```

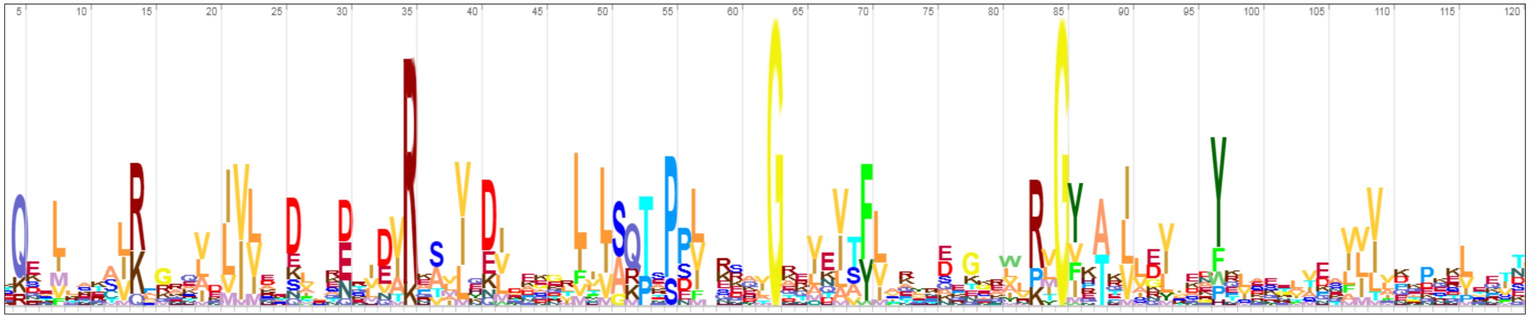


Figure S4. Sequence alignments of the newly identified xPilZ domains.

D. Sequence alignment of select representatives and an HMM logo of the PilZN6 domain, generated by the HMMer search at the EBI web site https://www.ebi.ac.uk/Tools/hmmer/. The top line shows secondary structure prediction by JPred. The two sequences from *Desulfarculus baarsii* are linked to their UniProt entries, sequences from other organisms are linked to the respective entries in NCBI protein database.

```
                                                    --EEE----EEEEE----HHHHHHHHHH-----EEEEEEHHHHHH-------EEEEEEE---EEEEEEEEEEEEE-----EEEHHHHHHHHH--
Cthe_0108_H.thermocellum             1   --MEIKVGDLVSIRHFSGTKLFKSLVLESN-NDEVLLKLTEDIALLNCSIGDPIVLGCELEENVYISSCTTLNVDKQQ-NTLKLKIDNYETTSN   90
Cthe_0697_H.thermocellum             5   LLPQNDSPIFCRVSINNDEKWVHGIITYIN2KNIAEIFLSAKYFKSYFNEGSKIVVKSLDENIETLFSGSVSKKVISI2QAITIQINKVLSYNN  100
Cthe_0733_H.thermocellum             1   --MKLRVGEIVTLKHYSAKRSNRGIISNSF-DNNVVIKPERDFLIYNFFDSDPVVIGFENDNVVNICESTVKHVDYNQ-NTFTLTINHIQSITN   90
Cthe_0868_H.thermocellum             6   LRDFLSEGAIIRTKHCNSTNWVTNVVYAIN2WIEVDIGLEKDYIDNIIMIGDTMRCKYSTDEYEFMLIGWVTKIKLEEPQSITIKIHDVERFAN  100
Cthe_0888_H.thermocellum            24   KEEIFRRGAVISIKHGYIFEPVLSVIQKCE-GECIYFRIPEEFLKNNVFKGDVVSCQVMQGEYEYIVSGIISEFEITYPWLVEVAIKRVSKVKN  116
Cthe_1065_H.thermocellum            21   PAKVVKLGSGVDIKHFKMFEYINAEVLGKT-DDSIKVEAKDVLKETLFFPGDHVVVNYSNTQDLFVMDGNINHVYDINPLIIDVKINKVEKLKD  113
Cthe_1173_H.thermocellum             1   ----MKVGEIASICHYSGKTWFKSVVLQVE-KDLLVVKLVKEFVILNFLENDPVVLGYEENDEVFIAGCIVKGIDPDC-SYIKLKVESVRSLKE   88
Cthe_1865_H.thermocellum             6   VSKVLKEGTVVNTKLKNGDIWIQNIVYRTE-KHMISIALLNEYLENIIMLGQTITIKYSSEHSEILFEGEIVKIRPEYPSYITIDIKDIKEVKN   98
Cthe_2893_H.thermocellum            11   LFNVMQYTNLVRVRFENDDVWTVLPILYAG-EDQVDILVDASKITLPGDVCGKAVIKFQRKGYEYIINGHIDNKSEDAPATVTIKYIDAKKYYN  103
RCX18322_Anaerobacterium             6   ISSILKEGLVASTKLGYDDTWHQNVIYRCV-ENIIWISLPQYYIQNIIMDGSKITIKYCNEFFEYLFEGVAFDIVPDYPAHVKVVVHSVHEIVN   98
RCX19951_Anaerobacterium            10   SEPIIPKGAVAFIRHQYADEPAMGIIRDIR-GKLIYICVNGDFKKYNPCKGEQVSCRLVDGRYEYDVCGSVEDFSAVNPAGICICVSSISKYEN  102
RCX08903_Anaerobacterium            11   LMNVMQDLNIVKVKFRNGDNWLIALLLYAD-ESNIEILFDSPYSENLLPY-QEAILKFQKEGYEYLVSGEVTTANASNPKAATLRVSMAQKYSD  102
RCX17134_Anaerobacterium             8   KEYNKSLSITALVFHYTVNLPCNAVVLSLS-EKNIVMRLEDESQQRDFSEGDPVVIAYNSGGEVIVKGFDVKKLNSDG--TIELDGDELESDAN   98
KNY24867_Pseudobacteroides           6   LGKYLKVGSIIKTKIDASPAWITNIVYSIS2FIQVNIGLEKNYIENLIMIGDTVKCKYTSNEIEIMFIGWVTKIHMDNPQRITIKVHQLNVFDN  100
KNY25286_Pseudobacteroides           3   NMEYFKRGTVISMSHASMSEPTLTVVQKIE-GHNLYFKLPMEFLKRNVFKGDKINAQVFGDEYELVINGIISDIDFKYPNYVQVCVDRIDKYRN   95
KNY25755_Pseudobacteroides          20   SLKVIKVSSDIGMKHYNVYNNLVATVELVN-DPIIKLTIKEGMKDLALAPRDIIVLSYNSDNDVYLLSASIISIEDDEPYQFTVSSLKIEKLKD  112
KNY26719_Pseudobacteroides          32   IINIMQKGSFIKTQFYGSNKWNVNLVLDNT-NNMLKVAFIEDDGIPALLPGDNIRCRFAVDTYDVNMLCTIDRMKCDFLPTKSLRILKADIWKN  124
KNY26764_Pseudobacteroides           6   VMAYLKPGYFVRVQASQNNLWSMNTFLKVN-RDSLVMPVTNELLNASVLINDFVKCKFNIGKNIITLTCMVEDISLALPQTIKLKIIKIDVFQD   98
KNY27276_Pseudobacteroides           7   IENILPEKSIIKTRHPGSYDWVTTIIRKAY-GNCIEVEQVEDYMVKVIMIGDNLLIKYSDNEYLYTIESTVQEIKFAS-RAVLLTVNSIKRIKN   98
KNY28482_Pseudobacteroides           1   --MNIKAGEIVSIRHCSGTKLYKSIVLDTI-DDILSVKLFEEVALLNCSPGDPVVLGCELDNEVYMGSCNLLDIDKDN-NKLSLKIDNYETITN   90
KNY29812_Pseudobacteroides           1   --MELKTDKIVLVSHYSAVKPFRCDIIDAS-DSSITLRLTKQFSILNFLEGDPAVIMIKEQNNIINIGCNVTSIEPKA-NVIKLRIDTIEPGSE   90
KNY30394_Pseudobacteroides           1   --MFIKPGEIISISHHSINILRSEVVDVQD4GQVLAIHIMKEFADCMLFEGDTVAIAFRSGDNIYTTSCHIINIDIDK-NVMQLVVENEEYVIN   94
OPX45492_Ruminiclostridium           6   ISKLIVPGSLVSFRKLNQSAWSLNIITECI-EDTICLPFTNDLLRACLFSGTPVDIKFNNQYYEYILNCSVVKIELSRFPYIRVKVHNICESQN   98
EMS70862_Ruminiclostridium           6   ISKLIIPGALISFRIFNQTVWSLNIVSQCS-GDTLNIPLTNDLMKACLFTNTNVEIKYKNEYFEYNIKGVISKIELSASPYINVKIDHVAANLN   98
```

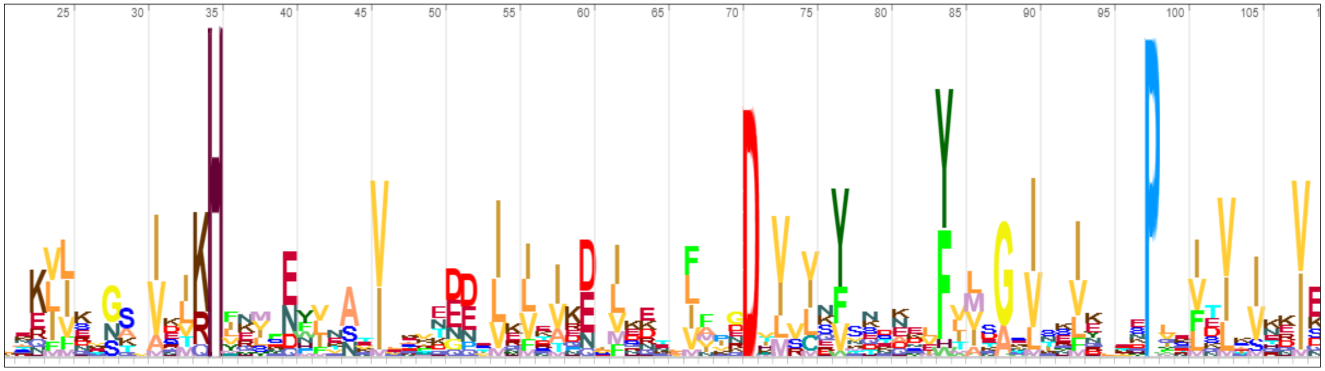


Figure S4. Sequence alignments of the newly identified xPilZ domains.

E. Sequence alignment of select representatives and an HMM logo of the PilZN7 domain, generated by the HMMer search at the EBI web site https://www.ebi.ac.uk/Tools/hmmer/. The top line shows secondary structure prediction by JPred. The sequences from *Hungateiclostridium* (formerly *Clostridium*) *thermocellum* are linked to UniProt, sequences from other organisms are linked to the respective entries in NCBI protein database.