

Supplemental Material

Table S1: BLASTP search results with phi92 proteins against the Enterobacteria phage rv5 genome (Genbank accession number: NC_011041.1).

Bold numbers indicate different positions (>50 amino acids) of aligned sequences in the two proteins. Phi92 proteins of which more than one stretch shares sequence similarity to more than one stretch of one or two rv5 proteins are *shaded in gray*.

Phi92 gp	Rv5 gp (Genbank accession number)	Aligned residues (phi92 gp rv5 gp)	% length alignment (phi92 gp)	Identity [%] E-value Positives [%] Bif score	Predicted function of the phi92 protein
004	206 (YP_002003508.1)	6-63 3-62	82.9%	50.0 65.0 2.0E-17 62	Conserved hypothetical protein
038	080 (YP_002003582.1)	1-268 1-251	99.3%	47.3 63.2 3.0E-89 258	Putative Sit2-like transferase
056	118 (YP_002003620.1)	11-169 15-169	94.1%	47.8 64.6 2.0E-44 138	Putative DNA-methylase/DNA-methyltransferase
060	115 (YP_002003711.1)	56-193 60-200	55.2%	27.1 45.1 2.0E-07 43	Putative phoH
063	107 (YP_002003609.1)	3-81 9-84	91.9%	39.2 57.0 1.0E-16 62	Putative glutaredoxin 1 (nrdC-like)
073	106 (YP_002003608.1)	1-356 1-358	98.6%	51.1 69.3 1.0E-126 361	Putative class I ribonucleotide reductase (RNR2 subunit) nrdB-like
075	105 (YP_002003607.1)	9-742 2-738	97.3%	53.3 73.4 0.0 833	Putative class I ribonucleotide reductase (RNR1 subunit) nrdA-like
077	103 (YP_002003715.1)	43-102 33-90	44.1%	35.0 55.0 9.0E-06 36	Conserved hypothetical protein
078	102 (YP_002003604.1)	1-320 1-323	99.1%	52.2 66.9 1.0E-114 328	Putative thymidylate synthase
079	100 (YP_002003602.1)	1-194 1-194	99.0%	33.2 50.0 2.0E-24 89	Conserved hypothetical protein
084	095 (YP_002003597.1)	1-320 2-326	100.0%	39.6 56.8 2.0E-76 230	Putative DNA polymerase (3'-5'-exonuclease domain)
085	090 (YP_002003592.1)	1-335 31-374	97.4%	39.4 60.3 3.0E-86 258	Putative 5'-exodeoxyribonuclease
089	081 (YP_002003583.1)	1-230 6-230	95.8%	30.2 45.0 2.0E-13 60	Conserved hypothetical protein
091	084 (YP_002003586.1)	1-371 1-421	100.0%	30.0 46.7 1.0E-51 170	Putative ATP dependent DNA ligase
095	077 (YP_002003579.1)	2-134 3-136	99.3%	42.5 64.2 2.0E-38 120	Putative polynucleotide kinase
096	078 (YP_002003580.1)	4-308 11-304	98.4%	46.9 65.3 1.0E-93 273	Putative RNA ligase / RNA repair
097	076 (YP_002003578.1)	2-161 4-170	87.4%	36.7 56.8 2.0E-34 113	Putative phosphoesterase
100	079 (YP_002003581.1)	4-110 5-113	94.7%	48.2 68.2 2.0E-30 98	Conserved hypothetical protein
103	088 (YP_002003590.1)	24-109 5-87	77.5%	46.5 66.3 4.0E-25 84	Putative transposase-like protein
114	069 (YP_002003571.1)	1-88 1-86	98.9%	29.2 52.8 2.0E-09 43	Conserved hypothetical protein
119	065 (YP_002003567.1)	195-669 22-507	68.4%	40.4 57.6 2.0E-107 329	Potential large terminase with N-terminal endonuclease domain
120	064 (YP_002003566.1)	5-522 6-518	99.2%	41.3 59.4 2.0E-126 373	Potential portal protein
121	063 (YP_002003565.1)	41-153 50-166	71.1%	41.5 61.9 1.0E-20 77	Conserved hypothetical protein
124	060 (YP_002003562.1)	9-331 10-331	97.0%	43.0 57.9 1.0E-75 229	Major capsid protein (MCP)
126	057 (YP_002003559.1)	7-164 42-201	86.8%	25.1 47.4 4.0E-07 41	Potential tail capping protein - T4 baseplate gp15 ortholog
127	056 (YP_002003558.1)	1-114 4-108	70.4%	30.7 49.1 4.0E-06 37	Conserved phage protein
128	055 (YP_002003734.1)	2-100 7-115	71.7%	27.5 52.3 8.0E-13 55	Conserved hypothetical protein
130	053 (YP_002003555.1)	4-463 3-453	98.3%	27.6 48.2 3.0E-49 167	Potential tail sheath
131	052 (YP_002003554.1)	10-156 10-156	92.5%	44.2 58.5 2.0E-33 109	Potential tail tube
132	051 (YP_002003553.1)	14-152 6-138	86.9%	29.3 50.7 8.0E-18 69	Conserved hypothetical protein
133	050 (YP_002003552.1)	1-57 53-108	76.0%	32.8 60.3 6.0E-07 37	Predicted ORF. No RBS identified. Bad Prodigal statistics.
135	048 (YP_002003759.1)	59-232 61-241	73.7%	26.4 44.0 1.0E-14 64	Potential tail tube initiator or completion protein
136	047 (YP_002003549.1)	39-131 18-112	66.9%	41.7 59.4 3.0E-19 71	Conserved hypothetical protein

137	046 (YP_002003548.1)	67-324 53-318	76.8%	38.6 57.3 1.0E-51 167	Potential baseplate hub
138	045 (YP_002003547.1)	8-234 8-223	92.7%	27.2 45.6 6.0E-23 86	Baseplate central spike
139	044 (YP_002003546.1)	6-179 7-176	82.5%	25.1 45.1 1.0E-14 63	Potential T4 baseplate gp25 ortholog
140	041 (YP_002003744.1)	1-37 1-37	23.6%	35.1 64.9 6.0E-06 38	Conserved phage structural protein
	043 (YP_002003545.1)	44-157 109-222	72.6%	41.7 61.7 7.0E-28 97	
141	042 (YP_002003544.1)	3-927 2-863	96.6%	35.2 52.1 7.0E-177 529	Potential tail fiber protein with glycosidase activity or carbohydrate binding module.
142	034 (YP_002003536.1)	1-686 1-684	95.4%	42.2 61.3 0.0 597	Potential tail fiber protein with glycosidase activity or carbohydrate binding module.
144	029 (YP_002003531.1)	8-88 3-83	92.0%	35.4 54.9 4.0E-16 60	Conserved phage structural protein
145	036 (YP_002003538.1)	1-492 1-493	99.4%	31.7 48.7 1.0E-65 213	Potential T4 baseplate gp6 ortholog
146	035 (YP_002003537.1)	5-204 8-200	95.7%	29.4 50.8 2.0E-27 97	Conserved phage structural protein
147	030 (YP_002003532.1)	1-324 1-345	99.4%	23.7 39.8 2.0E-19 81	Potential tail fiber or connector protein
	033 (YP_002003535.1)	1-323 1-343	99.1%	27.8 43.8 7.0E-31 112	
148	032 (YP_002003534.1)	6-181 3-192	96.2%	30.7 51.0 2.0E-24 88	Potential tail fiber chaperon
150	028 (YP_002003754.1)	10-114 6-112	13.1%	43.9 64.5 4.0E-23 97	Potential colanidase, degradation of colanic acid.
151	041 (YP_002003755.1)	23-91 52-120	10.6%	37.7 50.7 4.0E-08 49	Potential tail fiber protein
	028 (YP_002003756.1)	557-652 486-581	14.7%	48.5 66.0 3.0E-23 96	
153	027 (YP_002003757.1)	102-288 1218-1419	63.4%	26.3 40.6 4.0E-07 44	Conserved phage protein
		47-129 49-136	26.3%	43.2 59.1 5.0E-13 61	Conserved hypothetical protein
156	209 (YP_002003758.1)	161-315 252-423	49.2%	35.1 51.2 6.0E-21 85	
		92-572 40-614	43.6%	23.2 38.9 1.0E-21 94	
159	223 (YP_002003760.1)	843-966 635-759	11.2%	36.8 53.6 1.0E-19 87	Putative DNA polymerase
161	225 (YP_002003762.1)	113-583 67-618	79.8%	21.2 40.5 7.0E-12 60	Primase/helicase (T7-like)
170	002 (YP_002003504.1)	1-169 1-162	100.0%	39.7 55.8 2.0E-38 122	Putative cell wall hydrolase SleB
185	204 (YP_002003506.1)	88-236 6-150	62.6%	32.0 49.0 7.0E-16 66	Putative phosphatase
191	112 (YP_002003614.1)	1-153 1-153	98.1%	43.8 62.8 8.0E-46 140	Putative anaerobic ribonucleotide reductase (class III) activating protein (nrdG-like)
192		9-321 388-702	97.5%	64.7 79.5 1.0E-147 425	Putative (class III) anaerobic ribonucleotide reductase (nrdD-like), Part II
193	108 (YP_002003610.1)	4-286 99-387	98.3%	68.2 84.1 2.0E-147 423	Putative (class III) anaerobic ribonucleotide reductase (nrdD-like), Part I
192A		4-598 99-692	99.3%	77.5 87.2 0.0 979	Putative (class III) anaerobic ribonucleotide reductase (nrdD-like), Spliced variant
201	213 (YP_002003515.1)	1-120 1-121	99.2%	33.6 50.8 2.0E-13 56	Conserved hypothetical protein
212	198 (YP_002003500.1)	3-64 4-73	93.9%	41.4 64.3 1.0E-13 53	Conserved hypothetical protein
218	018 (YP_002003770.1)	15-81 16-78	80.7%	23.9 50.8 4.0E-06 34	Hypothetical protein
227	179 (YP_002003681.1)	2-119 5-122	98.3%	67.8 79.7 4.0E-59 171	Conserved hypothetical protein
247	170 (YP_002003672.1)	1-75 1-75	98.7%	90.7 97.3 3.0E-50 145	Conserved hypothetical protein

Table S2: BLASTP search results with phi92 proteins against the *Salmonella* phage PVP-SE1 genome (Genbank accession number: GU070616.1).

Bold numbers indicate different positions (>50 amino acids) of aligned sequences in the two proteins. Phi92 proteins of which more than one stretch shares sequence similarity to more than one stretch of one or two PVP-SE1 proteins are *shaded in gray*.¹ Note: the gp numbers of PVP-SE1 are according to the recent publication (1). In the genome Genbank file (NC_016071.1) the gps 015–144 are shifted by one (i.e. gene 16 encodes for gp015...).

Phi92_gp	PVP-SE1_gp (Genbank accession number) ¹	Aligned residues (phi92_gp PVP-SE1_gp)	% length alignment (phi92_gp)	Identity [%] Positives [%] E-value Bit score	Predicted function of the phi92 protein
038	109 (ADP02504.1)	1-268 1-247	99.3%	54.9 71.3 2.0E-110 312	Putative Sir2-like transferase
047	148 (ADP02544.1)	5-191 3-209	97.4%	39.1 51.7 6.0E-34 113	Conserved hypothetical protein
056	152 (ADP02548.1)	1-169 5-167	100.0%	47.1 65.9 4.0E-49 151	Putative DNA-methylase/DNA-methyltransferase
060	147 (ADP02543.1)	3-193 4-199	76.4%	26.3 43.4 1.0E-09 50	Putative phoH
063	141 (ADP02536.1)	3-81 8-83	91.9%	35.4 59.5 8.0E-15 57	Putative glutaredoxin 1 (nrdC-like)
073	139 (ADP02534.1)	1-361 1-362	100.0%	52.6 69.4 3.0E-134 380	Putative class I ribonucleotide reductase (RNR2 subunit) nrdB-like
074	031 (ADP02426.1)	33-203 20-195	78.4%	35.5 48.1 1.0E-26 95	Conserved domain protein
	105 (ADP02500.1)	32-178 36-179	67.4%	35.5 51.3 6.0E-23 86	
	204 (ADP02600.1)	45-170 44-166	57.8%	35.7 54.0 3.0E-24 88	
075	138 (ADP02533.1)	9-742 2-738	97.3%	53.8 74.4 0.0 837	Putative class I ribonucleotide reductase (RNR1 subunit) nrdA-like
077	136 (ADP02531.1)	39-105 29-100	49.3%	31.9 47.2 1.0E-06 39	Conserved hypothetical protein
078	134 (ADP02529.1)	1-320 1-321	99.1%	62.7 74.3 3.0E-146 408	Putative thymidylate synthase
079	131 (ADP02526.1)	1-173 5-179	88.3%	34.3 48.9 4.0E-25 90	Conserved hypothetical protein
084	128 (ADP02523.1)	12-320 7-323	96.6%	40.9 59.8 8.0E-84 249	Putative DNA polymerase (3'-5'-exonuclease domain)
085	123 (ADP02518.1)	1-335 26-368	97.4%	42.5 62.4 5.0E-95 280	Putative 5'-exodeoxyribonuclease
089	113 (ADP02508.1)	17-237 20-230	92.1%	37.2 50.7 3.0E-30 105	Conserved hypothetical protein
091	110 (ADP02505.1)	6-370 8-432	98.4%	29.3 46.1 8.0E-48 160	Putative ATP dependent DNA ligase
095	108 (ADP02503.1)	2-134 3-135	99.3%	67.2 81.3 1.0E-63 184	Putative polynucleotide kinase
096	106 (ADP02501.1)	4-308 3-307	98.4%	78.4 90.8 0.0 517	Putative RNA ligase / RNA repair
097	103 (ADP02498.1)	3-176 7-197	95.1%	38.9 52.9 1.0E-34 114	Putative phosphoesterase
100	107 (ADP02502.1)	4-113 2-114	97.3%	46.1 70.4 5.0E-30 97	Conserved hypothetical protein
103	102 (ADP02497.1)	24-101 8-85	70.3%	53.9 76.9 2.0E-32 102	Putative transposase-like protein
107	097 (ADP02492.1)	16-270 12-265	85.9%	40.9 60.2 3.0E-58 181	Putative phosphoribosyl pyrophosphate synthetase
108	096 (ADP02491.1)	9-563 11-560	96.2%	46.4 66.2 2.0E-169 486	Putative nicotinamide phosphoribosyltransferase
116	093 (ADP02488.1)	1-23 1-23	40.4%	60.9 82.6 2.0E-07 37	Conserved hypothetical protein
119	076 (ADP02471.1)	176-665 3-488	70.6%	39.5 57.5 2.0E-108 331	Potential large terminase with N-terminal endonuclease domain
120	075 (ADP02470.1)	42-522 28-516	92.1%	42.1 62.2 2.0E-130 383	Potential portal protein
121	074 (ADP02469.1)	36-154 63-185	74.8%	37.9 55.7 2.0E-14 61	Conserved hypothetical protein
123	072 (ADP02467.1)	14-125 9-126	86.2%	34.8 51.7 3.0E-08 42	Potential decoration protein
124	071 (ADP02466.1)	8-331 10-333	97.3%	40.4 57.8 5.0E-75 228	Major capsid protein (MCP)
126	065 (ADP02460.1)	5-149 4-144	79.7%	28.7 47.8 5.0E-09 46	Potential tail capping protein - T4 baseplate gp15 ortholog
127	064 (ADP02459.1)	1-93 1-89	57.4%	33.0 48.9 2.0E-07 41	Conserved phage protein
128	063 (ADP02458.1)	2-100 6-115	71.7%	29.1 53.5 1.0E-12 54	Conserved hypothetical protein
130	061 (ADP02456.1)	4-468 3-468	99.4%	27.8 47.7 7.0E-51 172	Potential tail sheath
131	060 (ADP02455.1)	10-153 10-156	90.6%	44.2 57.8 1.0E-32 107	Potential tail tube

132	059 (ADP02454.1)	14-147 4-132	83.8%	34.8 54.1 4.0E-20 75	Conserved hypothetical protein
133	058 (ADP02453.1)	1-62 18-81	82.7%	39.1 64.1 2.0E-12 50	<i>Predicted ORF. No RBS identified. Bad Prodigal statistics.</i>
135	056 (ADP02451.1)	28-208 25-219	76.7%	30.3 45.1 3.0E-20 80	Potential tail tube initiator or completion protein
136	055 (ADP02450.1)	30-131 6-105	73.4%	35.3 57.8 3.0E-22 79	Conserved hypothetical protein
137	054 (ADP02449.1)	17-324 7-317	91.7%	35.8 53.7 6.0E-59 186	Potential baseplate hub
138	053 (ADP02448.1)	9-235 10-225	92.7%	26.8 44.4 6.0E-17 70	Baseplate central spike
139	052 (ADP02447.1)	6-175 7-172	80.6%	31.0 43.3 2.0E-16 68	Potential T4 baseplate gp25 ortholog
140	049 (ADP02444.1)	1-48 1-49	30.6%	44.9 61.2 6.0E-07 40	Conserved phage structural protein
	051 (ADP02446.1)	1-91 1-95	58.0%	33.3 52.1 5.0E-08 44	
141	050 (ADP02445.1)	55-156 122-223	65.0%	48.0 67.7 6.0E-33 109	Potential tail fiber protein with glycosidase activity or carbohydrate binding module.
144	042 (ADP02437.1)	1-927 1-927	96.8%	77.6 89.0 0.0 1569	Conserved phage structural protein
145	048 (ADP02443.1)	2-492 3-494	99.2%	31.4 49.8 3.0E-68 220	Potential T4 baseplate gp6 ortholog
146	047 (ADP02442.1)	5-191 8-187	89.5%	31.4 51.6 8.0E-27 96	Conserved phage structural protein
147	046 (ADP02441.1)	1-77 1-78	23.6%	35.9 53.9 5.0E-11 56	Potential tail fiber or connector protein
		258-326 304-372	21.2%	65.2 79.7 1.0E-26 101	
148	045 (ADP02440.1)	6-183 4-179	97.3%	45.5 61.2 2.0E-49 151	Potential tail fiber chaperon
150	041 (ADP02436.1)	10-114 5-111	13.1%	49.5 64.2 2.0E-23 97	Potential colaminidase, degradation of colanic acid.
	049 (ADP02444.1)	208-798 113-708	73.5%	37.5 53.9 3.0E-119 369	
	051 (ADP02446.1)	4-74 34-104	10.9%	38.0 54.9 1.0E-11 57	
	049 (ADP02444.1)	11-73 39-102	9.6%	41.5 61.5 8.0E-09 51	
151		131-584 420-881	69.5%	28.2 40.8 2.0E-29 116	
	069 (ADP02464.1)	127-527 187-637	61.4%	28.9 42.0 2.0E-33 129	
		205-482 120-404	42.6%	27.6 37.7 7.0E-14 67	
		317-479 116-293	25.0%	29.7 45.1 2.0E-09 53	Potential tail fiber protein
152	041 (ADP02436.1)	529-652 481-601	19.0%	42.9 58.7 1.0E-20 88	
156	040 (ADP02435.1)	17-185 396-582	83.7%	39.6 53.5 4.0E-28 103	Putative tail component (lambda protein V like)
159	039 (ADP02434.1)	1-122 1-134	38.7%	32.6 48.6 4.0E-11 56	Conserved hypothetical protein
		161-315 249-409	49.2%	37.6 52.7 2.0E-26 100	
161	016 (ADP02411.1)	111-574 75-628	42.0%	24.0 38.1 2.0E-23 99	Putative DNA polymerase
162	014 (ADP02409.1)	108-499 62-530	66.4%	36.9 52.9 4.0E-22 95	Primase/helicase (T7-like)
167	008 (ADP02403.1)	5-89 10-93	95.5%	32.2 56.3 1.0E-09 44	Conserved hypothetical protein
170	006 (ADP02401.1)	13-62 5-54	78.1%	40.0 62.0 5.0E-08 38	Conserved hypothetical protein
185	243 (ADP02639.1)	8-169 9-172	95.9%	37.1 58.7 8.0E-37 119	Putative cell wall hydrolase SleB
191	232 (ADP02628.1)	90-237 7-159	62.2%	45.2 64.5 2.0E-40 130	Putative phosphatase
192	143 (ADP02538.1)	1-153 1-153	98.1%	66.7 85.0 5.0E-83 234	Putative anaerobic ribonucleotide reductase (class III) activating protein (nrdG-like)
193	142 (ADP02537.1)	9-320 382-692	97.2%	76.0 85.9 1.0E-178 504	Putative (class III) anaerobic ribonucleotide reductase (nrdD-like), Part II
192A		4-286 99-381	98.3%	79.2 88.7 3.0E-169 478	Putative (class III) anaerobic ribonucleotide reductase (nrdD-like), Part I
202	029 (ADP02424.1)	4-598 99-692	99.0%	77.5 87.2 0.0 979	Putative (class III) anaerobic ribonucleotide reductase (nrdD-like). Spliced variant
205	175 (ADP02571.1)	28-109 1-83	60.3%	53.0 72.3 7.0E-20 72	Putative NrdA, 1-like ribonucleoside-diphosphate reductase 1, alpha subunit
		4-45 3-49	26.3%	44.7 61.7 4.0E-08 41	Conserved hypothetical protein

217	226 (ADP02622.1)	6-65 4-63	32.8%	45.0 70.0 2.0E-15 61	Conserved hypothetical protein
218	100 (ADP02495.1)	3-77 18-92	90.4%	26.7 49.3 7.0E-06 34	Hypothetical protein
223	121 (ADP02516.1)	1-169 1-161	92.3%	43.9 56.7 4.0E-37 120	Conserved hypothetical protein
231	197 (ADP02593.1)	35-121 26-111	71.9%	30.9 46.8 4.0E-09 44	Conserved hypothetical protein

REFERENCES

1. Santos SB, Kropinski AM, Ceyssens PJ, Ackermann HW, Villegas A, Lavigne R, Krylov VN, Carvalho CM, Ferreira EC, Azeredo J. 2011. Genomic and proteomic characterization of the broad host range Salmonella phage PVP-SE1 - The creation of a new phage genus. *J. Virol.*

Figures

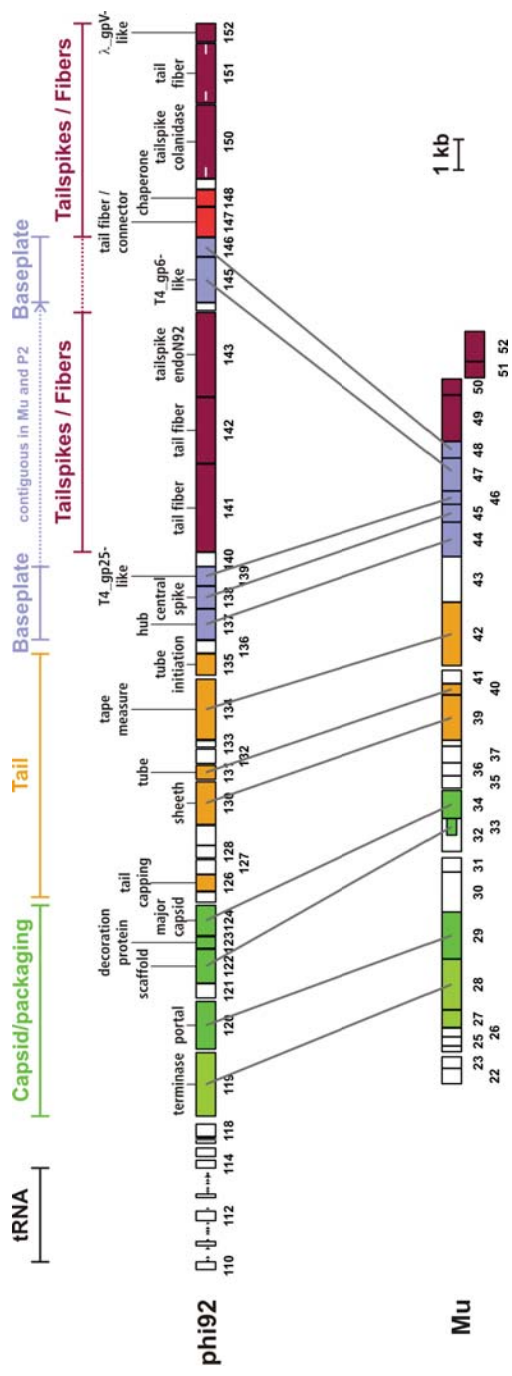


Figure S1: Comparison of structural operons of phages phi92 and Mu. *A*, the structural operons are drawn at the same scale (the *bar* in the right bottom corner represents 1 kb length). Genes encoding a protein are represented by *boxes* filled with the respective *color* according to the predicted function. tRNA genes are given as *black arrows*. Gene numbers are given below the boxes. Note that genes 51 and 52 of phage Mu are encoded on the complementary strand. The genes can be grouped according to the location and function of their products in the phage virion (*bars* above) represented in *different colors*. Functions assigned to phi92 proteins are given above the phi92 genome. *Gray solid lines* between two organisms represent homolog proteins.

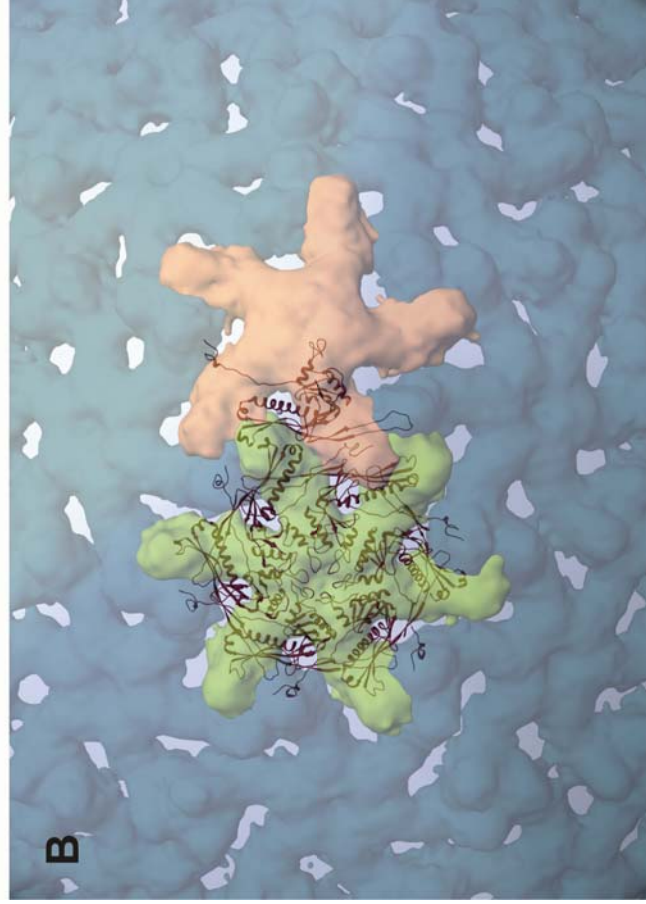
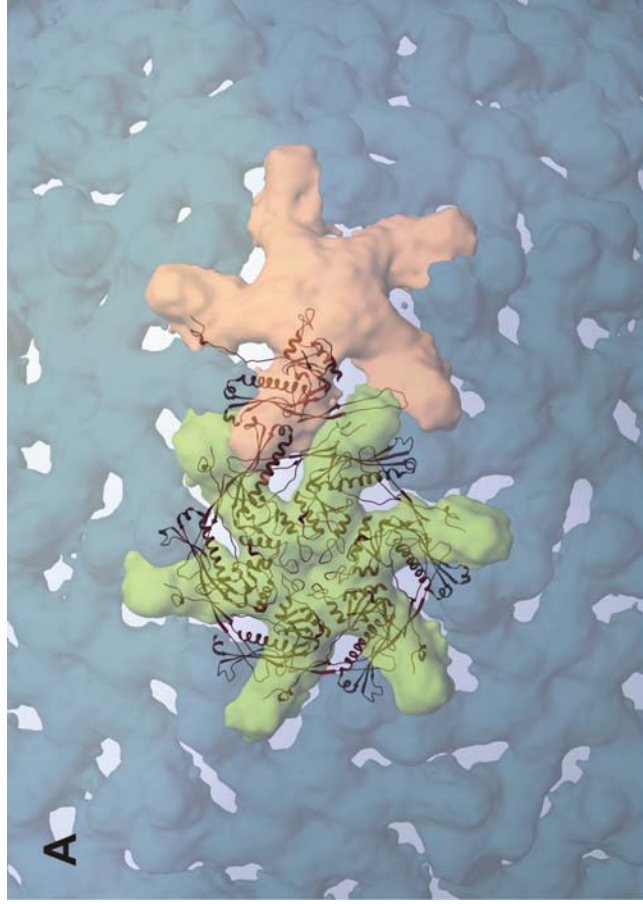


Figure S2: Hand determination of the phi92 capsid reconstruction. *A*, the original map (correct hand) and the best fit of the HK97 asymmetric unit structure. *B*, the mirrored map (wrong hand) and the best fit of the HK97 asymmetric unit structure. Both cryoEM maps are contoured at 2.5 σ .