Supplementary material for: *In vivo* self-assembly of stable GFP fusion protein particles and their uses in enzyme immobilization.

Mark Venning-Slater, David O. Hooks, Bernd H. A. Rehm

Institute of Fundamental Sciences, Massey University, Private Bag 11222, Palmerston North, New Zealand; telephone: +64-6-350-5515, ext. 7890; fax: +64-6-350-2267; e-mail: b.rehm@massey.ac.nz.

Table S1: Bacterial strains, plasmids, and primers used in this study.

Name	*Characteristics*	*Source*
<u>Bacterial strains</u>		
XLI-Blue	*recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*(r^-_k, m^+_k), *supE44*, *relA1*, -, *lac*[*F'*, *proAB*, *lacI*^{q}, *lacZΔM15*, *Tn10*(*Tc*^{r})]	Stratagene
BL21(DE3)	F-, *ompT*, *hsdS*$_B$ (r_B- m_B-), *gal*, *dcm* (DE3)	Novagen
<u>Plasmids</u>		
pMCS69	pBBR1MCS derivative containing genes *phaA* and *phaB* of *R. eutropha* colinear to lac promoter, Cp^R	(Amara & Rehm, 2003)
pUC57:nanA	Ap^r, containing *nanA* on the pUC57 plasmid	Genscript Corporation
pET14b	Ap^r, T7 promoter	Novagen
pET14b-nanA-phaC	pET14b vector containing *nanA* at the 3' end of *phaC*	(Hooks et el., 2013)
pET14b-ext(AVTS)gfp-phaC(C319A)-linker-zz	pET14b vector containing a gene fusion of ext(AVTS)gfp-phaC(C319A)-linker-zz	(Jahns et al., 2013)
pET14b-ext(AVTS)gfp-phaC(C319A)-linker-nanA	pET14b-ext(AVTS)gfp-phaC(C319A)-linker-zz derivative containing *nanA* fragment from pGEM-nanA at the 3' end	This study
pET14b-ext(AVTS)gfp-phaC(C319A)-linker-bla(-ss)	pET14b-ext(AVTS)gfp-phaC(C319A)-linker-zz derivative containing *bla(-ss)* fragment from pGEM-bla(-ss) at the 3' end	This study
pET14b-ext(AVTS)gfp-phaC(C319A)-linker-opdA	pET14b-ext(AVTS)gfp-phaC(C319A)-linker-zz derivative containing *opdA* fragment from pET14b-phaC-linker-opdA at the 3' end	This study
pET14b-ext(AVTS)gfp-nanA-linker-zz	pET14b-ext(AVTS)gfp-phaC-linker-zz derivative containing *nanA* in the *Nde*I site	This study
pET14b-ext(AVTS)gfp-nanA-linker-nanA	pET14b-ext(AVTS)gfp-nanA-linker-zz derivative containing *nanA* fragment from pGEM-nanA at the 3' end	This study
pET14b-bla(-ss)-phaC	pET14b vector containing a gene fusion of bla(-ss)-phaC without the signal sequence encoding region	(Rasiah & Rehm, 2009)
pET14b-phaC-linker-opdA	pET14b vector containing a gene fusion of phaC-linker-opdA	(Blatchford et al., 2012)
<u>Primers</u>		
Bla(-ss)-Fwd	ATA ATA CTC GAG ATG GCC GCT AAC CTG AAC GG	IDT
Bla(-ss)-Rev	AAT AAT GGA TCC TTA GCC ACC ACC ACC ACC GC	IDT

MALDI-TOF protein identification

Table S2: Tryptic peptides of GFP fusion proteins as identified by MALDI-TOF MS .

Total bead protein was electrophoresed onto a gel and a band corresponding to the theoretical molecular weight of the fusion protein was isolated and analysed by mass spectrometry.

Fusion Protein (coverage)	Identified peptides (modifications)
GFP-iPhaC-L-NanA (27%)	F32-K46, S91-R101, F119-R127, V269-R284, I311-R322, F363-R373, A374-K387, F394-R417, L418-R428, I443-R455, Y498-R511, H512-R526, N527-R547 (carbamidomethyl), E619-R634 (carbamidomethyl), L704-R726 (carbamidomethyl), E727-K746, F749-K766, A812-R821, G862-K879, T1112-R1126 (carbamidomethyl, oxidation), A1142-R1151 (oxidation)
GFP-NanA-L-NanA (28%)	F32-K46, S91-R101, F119-R127, H174-K214, D221-K243 (oxidation), L278-R307, A372-K394, L395-K413, Q414-R426, T505-R519 (carbamidomethyl), K520-K528, A535-R544, G545-R556, G557-R569
GFP-iPhaC-L-Bla (38%)	G8-K30, L145-K160, V268-R283, I310-R321, D325-K337, F362-R372, F393-R416, I442-R454, N455-K481, Y497-R510, H511-R525, N526-R546, E618-R633 (carbamidomethyl), G634-K660, L703-R725 (carbamidomethyl), E726-K745, F748-K765, A811-R820, L881-K903, G904-K926, W1011-R1025, W1071-K1090, D1099-R1105, T1133-R1161 (two deamidated), E1211-K1226, K1248-R1269, Q1249-R1269, Q1299-R1312, S1313-R1339
GFP-iPhaC-L-OpdA (39%)	G8-K30, L145-K160, D220-K242 V268-R283, F362-R372, F393-R416, L417-R427, I442-R454, N455-K481, Y497-R510, H511-R525, N526-R546, E618-R633 (carbamidomethyl), L703-R725 (carbamidomethyl), E726-K745, A811-R820, G871-R896 (carbamidomethyl), A897-R905, A921-R937, A948-R968, S971-R981, A1005-K1014, D1037-R1054, V1055-R1075 (carbamidomethyl), M1084-R1104, I1149-R1160, G1169-R1185

Table S3: Long-term enzyme particle stability results.

Enzymes activities after storage for NanA and Bla enzymes are expressed as percentages of initial activity, while those of OpdA are in the original absorbance units. Storage periods are in brackets (days). n.d. = not detected, n.a. = not assayed, t_0 = initial production level before storage.

	Production level of Neu5AC (NanA), maltose (BLA), or *para*-nitrophenol (OpdA) after storage				
		Temperature			
Fusion protein	t_0	37 °C	RT	4 °C	-80 °C
GiCLN	50.6 ± 1.9	n.a.	n.d. (97)	45.1 (97)	48.4 (97)
NanA-PhaC	163.2 ± 1.9	n.a.	3.2 (97)	74.4 (97)	77.1 (97)
GiCLB	0.96 ± 0.04	n.a.	91.7 (95)	104.1 (95)	93.8 (69)
Bla(-ss)-PhaC	2.08 ± 0.01	n.a.	65.9 (95)	98.6 (95)	99.5 (69)
GiCLO	n.a	n.a.	0.52 ± 0.01 (132)	0.58 ± 0.01 (132)	0.65 ± 0.01 (132)
PhaC-OpdA	n.a.	n.a.	n.a.	1.01 ± 0.01 (133)	n.a.
GiCLO (2)	0.258± 0.002*	0.159 ± 0.003 (30)	n.a	0.166 (30)	0.166 (30)
PhaC-OpdA (2)	0.247± 0.005*	n.d.	n.a	0.157 ± 0.005 (30)	0.160 ± 0.002 (30)

*Measured in a separate assay to the storage results meaning initial substrate concentration could vary and are therefore not directly comparable, see Materials and Methods.

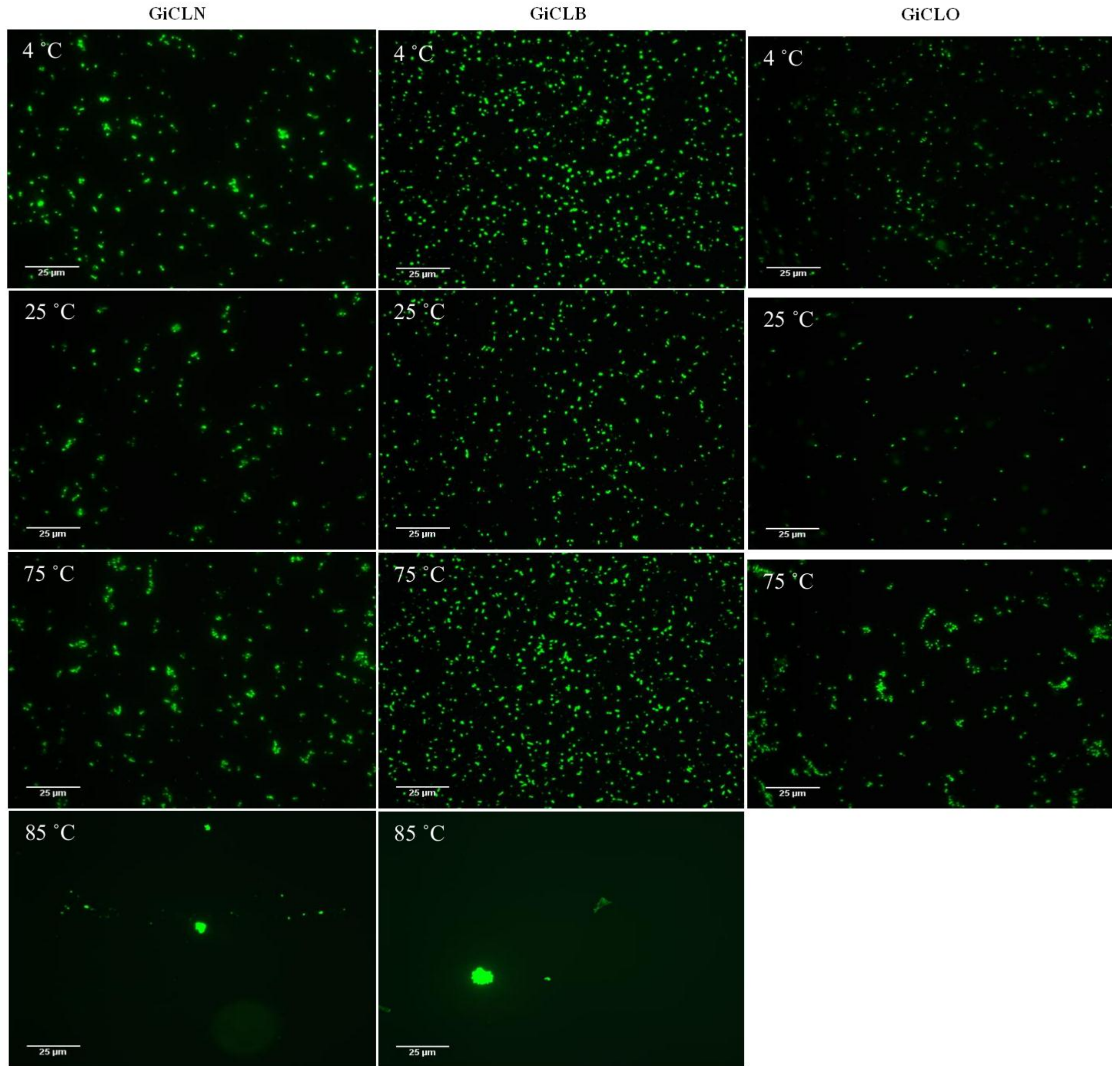

Figure S1: Fluorescent images of GFP particles after temperature treatment at 4 °C, 25 °C, 75 °C, and 85 °C were obtained at 1000x magnification using a U-MWIBA2 Blue excitation filter cube fitted to an Olympus BX51 Fluorescent Light Microscope (Olympus Optical Co., Japan), an Optronics camera (Optronics, USA), and MagnaFire™ 2.1C Application software (Optronics, USA). GiCLN: GFP-iPhaC-L-NanA, GiCLB: GFP-iPhaC-L-Bla, GiCLO: GFP-iPhaC-L-OpdA.

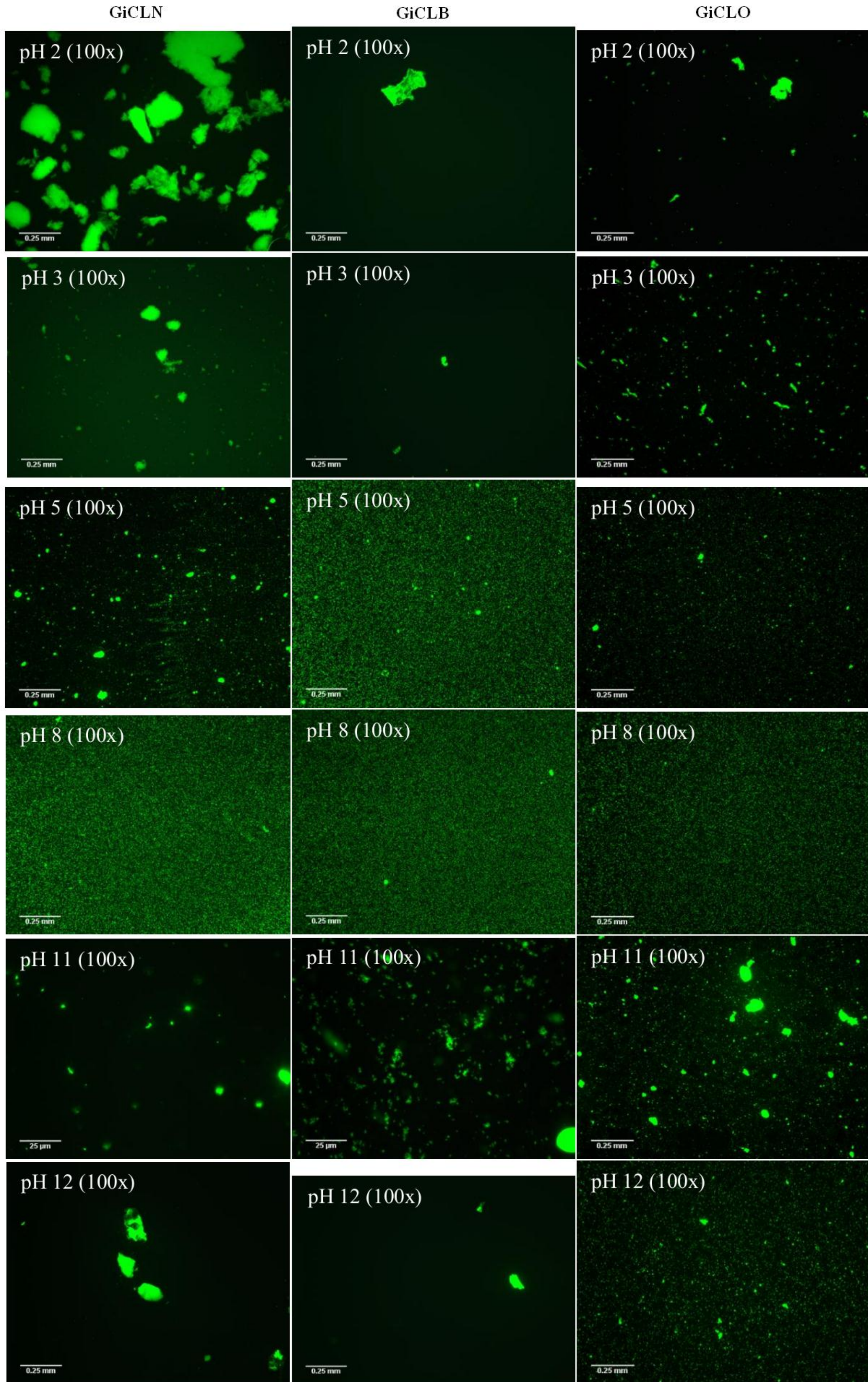

Figure S2: Fluorescent images of GFP particles after 10 min pH treatment at 2, 3, 5, 8, 11, and 12 were obtained at 100x magnification using a U-MWIBA2 Blue excitation filter cube fitted to an Olympus BX51 Fluorescent Light Microscope (Olympus Optical Co., Japan), an Optronics camera (Optronics, USA), and MagnaFire™ 2.1C Application software (Optronics, USA). GiCLN: GFP-iPhaC-L-NanA, GiCLB: GFP-iPhaC-L-Bla, GiCLO: GFP-iPhaC-L-OpdA.

References

1. Amara, A., Rehm, B. H.A (2003) Replacement of the catalytic nucleophile cysteine-296 by serine in class II polyhydroxyalkanoate synthase from Pseudomonas aeruginosa mediated synthesis of a new polyoxoester: Identification of catalytic residues. Biochem. J. 374:413-421.

2. Hooks, D.O., Blatchford, P.A., Rehm, B.H.A. (2013) Bioengineering of bacterial polymer inclusions catalyzing the synthesis of N-acetyl neuraminic acid. Appl. Environ. Microbiol. 79:3116-21.

3. Jahns, A.C., Maspolim, Y., Chen, S., Guthrie, J.M., Blackwell, L.F., Rehm, B.H.A. (2013) In vivo self-assembly of fluorescent protein microparticles displaying specific binding domains. Bioconjug. Chem. 24 (8), 1314-1323.

4. Rasiah, I, Rehm, B.H.A. (2009) One-step production of immobilised ⍰-amylase in recombinant Escherichia coli. Appl. Environ. Microbiol. 75: 2012–2016.

5. Blatchford, P.A., Scott, C., French, N., Rehm, B.H.A. (2012) Immobilization of organophosphohydrolase OpdA from Agrobacterium radiobacter by overproduction at the surface of polyester inclusions inside engineered Escherichia coli. Biotech. Bioeng. 109:1101-1108.