Ross, J. F., Bridges, A., Fletcher, J. M., Shoemark, D., Alibhai, D., Bray, H. E. V., ... Woolfson, D. N. (2017). Decorating Self-Assembled Peptide Cages with Proteins. *ACS Nano*, *11*(8), 7901-7914. https://doi.org/10.1021/acsnano.7b02368

Decorating Self-Assembled Peptide Cages with Proteins

*James F. Ross,[1] Angela Bridges,[2] Jordan M. Fletcher,[1] Deborah Shoemark,[3,4] Dominic Alibhai,[5] Harriet E. V. Bray,[1] Joseph L. Beesley,[1] William M. Dawson,[1] Lorna R. Hodgson,[4] Judith Mantell,[5] Paul Verkade,[5] Colin M. Edge,[2] Richard B. Sessions,[3,4] David Tew[2] and Derek N. Woolfson[1,3,4]**

[1]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

[2] GSK, Gunnels Wood Rd, Stevenage, SG21 2NY, UK

[3]BrisSynBio, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK

[4]School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD, UK

[5]Wolfson Bioimaging Facility, University of Bristol, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD, UK

*Address correspondence to d.n.woolfson@bristol.ac.uk or david.tew@gsk.com

Keywords: coiled coil; nanoreactor; protein cage; protein design; self-assembly; supramolecular assembly; synthetic biology.

ASBSTRACT

An ability to organize and encapsulate multiple active proteins into defined objects and spaces at the nanoscale has potential applications in biotechnology, nanotechnology and synthetic biology. Previously, we have described the design, assembly and characterization of peptide-based self-assembled cages (SAGEs). These ≈100 nm particles comprise thousands of copies of *de novo* designed peptide-based hubs that array into a hexagonal network and close to give caged structures. Here, we show that, when fused to the designed peptides, various natural proteins can be co-assembled into SAGE particles. We call these constructs pSAGE for protein-SAGE. These particles tolerate the incorporation of multiple copies of folded proteins fused to either the *N* or the *C* termini of the hubs, which modeling indicates form the external and internal surfaces of the particles, respectively. Up to 15% of the hubs can be functionalized without compromising the integrity of the pSAGEs. This corresponds to hundreds of copies giving mM local concentrations of protein in the particles. Moreover, and illustrating the modularity of the SAGE system, we show that multiple different proteins can be assembled simultaneously into the same particle. As the peptide—protein fusions are made *via* recombinant expression of synthetic genes, we envisage that pSAGE systems could be developed modularly to actively encapsulate or to present a wide variety of functional proteins, allowing them to be developed as nanoreactors through the immobilization of enzyme cascades, or as vehicles for presenting whole antigenic proteins as synthetic vaccine platforms.

In nature, the co-location of multiple bioactive molecules within defined nanoscale spaces underpins many biological functions. For example, in eukaryotes various sub-cellular organelles organize and separate different biochemical processes and cellular functions.[1-4] Sub-cellular organization also occurs in bacteria. For instance, bacterial micro-compartments (BMCs) allow the co-location of enzyme cascades and the isolation of potentially toxic intermediates, leading to enhanced product turnover. BMCs, certain viral capsids, and other protein containers are increasingly being adapted for biotechnology in applications such as drug delivery, antigen presentation and as enzymatic nanoreactors.[5] It has been shown that BMCs can be modified to incorporate completely different protein systems, including fluorescent proteins or enzyme cascades, either through protein-fusion strategies[6] or by using peptide tags.[7,8] Similarly, smaller protein assemblies, such as lumazine synthase can be engineered to encapsulate fluorescent proteins[9] and enzymes.[10]

In addition to these protein-engineering approaches, others are exploring rational design and assembly of addressable, nanoscale, biologically inspired containers.[11-13] This includes containers constructed from virus-like particles,[14,15] DNA origami,[16-18] vesicles,[19] natural[20-21] and *de novo* protein assemblies[22-27] and peptide shells.[28-30] More specifically, several groups have constructed protein-based supramolecular assemblies by exploiting the symmetry of natural multimeric proteins. This is achieved by fusing together two naturally occurring protein domains with different oligomerization properties, *e.g.*, dimer and trimer.[22] In this way, oligomerization of each of the linked domains propagates the assembly of the protein network, which, given the correct linker and subunit geometry, can close to form specified supramolecular objects.[24] In this way, supramolecular protein assemblies have been described for tetrahedral, octahedral, and icosahedral protein cages[23-24, 31] and for porous protein cubes.[25] Many of these designs have

been achieved using computational design methods such as Rosetta, including more recent two-component systems,[26, 32] which give more control over assembly.

Turning to peptide-based assemblies, Burkhard and colleagues describe the assembly of polyhedral nanoparticles[28] from linear peptides encoding both pentameric and trimeric coiled-coil oligomerization domains. More recently, we have used completely *de novo* designed peptides to assemble a two-component system, which we call self-assembled peptide-based cages (SAGEs), that forms spherical particles ≈100 nm in diameter.[29] These are made from ≈1500 copies of two types of peptide hub (HubA and HubB). Each hub comprises a homo-trimeric coiled coil (CC-Tri3)[33] each peptide chain of which is linked back-to-back, *via* a disulfide bond, to one half of a hetero-dimeric coiled coil.[34] The heterodimer has an acidic helix (CC-DiA) and a basic helix (CC-DiB), which do not fold on their own. Therefore, each hub effectively has a folded trimeric core, with 3 appended unfolded peptides. Our working model of how these co-assemble is as follows: when mixed, the CC-DiA and CC-DiB components of the hubs associate; the resulting combination of 2-fold and 3-fold axes of symmetry lead to a tessellated hexagonal peptide lattice, with ≈6 nm pores; and, finally, this network folds into the observed spherical particles. The main features of this model have been confirmed by scanning electron microscopy (SEM) and atomic force microscopy (AFM).[29] In addition, extensive molecular dynamics (MD) studies suggest that patches of the SAGE skin curve to leave the *N* termini of the CC-Tri units on convex faces of the patches and, therefore, the outer surfaces of SAGE particles.

Herein, we describe the incorporation of different protein fusions into SAGE particles, including Green Fluorescent Protein (GFP), mCherry (mCh), Maltose Binding Protein (MBP), and *Renilla* luciferase (Luc). We achieve this by making synthetic genes for these proteins plus

the CC-Tri units, and expressing the resulting protein fusions. Based on our working model of SAGE structure from the MD studies, we rationalized that fusing proteins to the *N* terminus of CC-Tri would present proteins predominantly on the outer surfaces of the SAGEs, whereas *N*-terminal fusions would be encapsulated within the particles. The fusion proteins are made into hubs again through disulfide bonds to the half-heterodimer units to give protein-HubA and protein-HubB, and these are mixed to generate protein-SAGE particles (pSAGEs). We show that the protein fusions fully incorporate into the SAGE particles, and that up to 15% of the hubs are addressable without compromising the integrity of either the SAGEs or the activity of the proteins. This approximates to mM effective local concentrations of active protein fusions. In addition, we show that multiple proteins can be incorporated when expressed as single fusion proteins, or co-assembled into pSAGEs as separate, multiple, fusion proteins.

Results and Discussion

Designing modules for assembling pSAGEs

To test the pSAGE concept, we generated a number of protein fusions containing the CC-Tri3 peptide sequence. Two of these fusions contained a cysteine-free GFP[35] fused to the *N* or *C* terminus of the CC-Tri3 sequence *via* a flexible 36-residue serine-glutamate-glycine-based (SEG) linker (Biobrick Part:BBa_K243030, Freiburg iGEM Team 2009). This gave GFP-CC-Tri3 and CC-Tri3-GFP constructs, respectively, Figure 1a,d. A third GFP fusion protein was made that lacked the CC-Tri3 domain but had a truncated *N*-terminal SEG tag to serve as a control (GFP-control). Each fusion protein was expressed from a synthetic gene in *E. coli*, and then purified *via* Ni-chelation chromatography utilizing *N- or C*-terminal His-tags included in each construct. Pure fusion proteins were covalently linked to the CC-Di-A peptide *via* an

asymmetric disulfide bond between the single cysteine residues harbored within the CC-Di-A and CC-Tri3 modules. The resulting assemblies are referred to as GFP-HubA and HubA-GFP, respectively, Figure 1b,d. pSAGE particles were then assembled by mixing equimolar amounts of HubA (combining parent-HubA and the protein-fusion variants) and K4-HubB (which has an *N*-terminal tetra-lysine tag on the CC-Tri3 component) to give GFP-SAGE and SAGE-GFP, respectively, Figure 1c. We used the K4 variant (Figure S2) as this has a *N*-terminal tetralysine tag that we find improves solubility of SAGE particles, presumably because of the increased positive charge. Typically, these mixtures were incubated at 20˚C for 1 hour prior to experiments. Additional pSAGE variants were constructed containing MBP (M in Figure 1a,b,d) and mCh or Luc. Examples of specific fusion proteins are given in Figure 1d, and a schematic of all protein fusions used in the study can be found in the supplementary information, Figure S1. *N.b.,* hubs and SAGE particles assembled without fusion proteins are referred to as 'parent-SAGE' and 'parent-Hubs, respectively.

Figure 1. Schematics for SAGE modules, nomenclature and assembly. (a) The homotrimer (CC-Tri3, green, with and without the tetra-lysine tag), heterodimer (CC-Di-A (red) and CC-Di-B (blue)) and a generic fusion protein that contains the CC-Tri3 module (CC-Tri3-Protein). This fusion protein can contain an *N*-terminal MBP (M) and/or a fusion protein to the *N*- or *C*-terminal side of the CC-Tri3 motif (1 and 2 respectively). **(b)** SAGE components: parent-HubA, K4-HubB and HubA-Protein each generated by disulfide-bond formation between the CC-Tri3 domain and respective CC-Di peptide. **(c)** Representations of: the hexagonal lattice formed upon mixing equimolar concentrations of parent-HubA and K4-HubB components; and a pSAGE particle formed with pendant HubA-Proteins. **(d)** Examples of fusion proteins with a schematic depicting the orientation of the proteins relative to the CC-Tri3 motif.

Heteromeric hubs comprising peptide and protein-fusions can be generated

First, to test the impact of introducing large proteins into the peptide-based modules and hubs, we measured the stabilities of one of the GFP-containing CC-Tri3 fusions by thermal-unfolding

experiments following the loss of α-helical signal of the CC-Tri3 unit in circular-dichroism (CD) spectroscopy. Compared with parent CC-Tri3, which has a midpoint unfolding temperature (T_M) of 56˚C at 50 μM (Figures S3c, d and e), the stability of CC-Tri3-GFP was reduced (T_M = 44˚C, Figures S2a, d and e). Consistent with this, the CC-Tri3-GFP fusion sediments as a species intermediate between dimer and trimer in analytical ultracentrifugation, Figure S3f. To test if this stability could be recovered, we mixed CC-Tri3 and CC-Tri3-GFP in a 2:1 ratio. This gave a T_M of 57˚C, Figure S3b, d and e.

These data indicate that the CC-Tri3 motif of CC-Tri3-GFP is folded at ambient temperature (Figure S3e), but its thermal stability is compromised by 12˚C at 50 μM protein. Presumably this is because of the three bulky GFP molecules being brought into close proximity. However, mixing the GFP construct with an excess of the CC-Tri3 peptide restores the stability of the CC-Tri3 motif, presumably because heterotrimers are formed.

Moving onto the hubs, size-exclusion chromatography (SEC) was used to probe the oligomeric states of HubA-GFP in solution. This gave peaks corresponding to monomers, dimers and trimers of the fusion protein, Figure S4b. When free HubA was added to HubA-GFP, the area of the "monomeric" peak, which in this experiment effectively reflects a single copy of the GFP fusion, increased at the expense of the other peaks, compare Figure S4b,c with Figure S4d,e. This suggests that the HubA-GFP protomers exchange and associate with the HubA peptide.

Again, these data are consistent with the pendant GFP destabilizing homotrimer formation, but that HubA-GFP and free HubA can combine *via* their common trimeric, coiled-coil units to give more-stable heteromeric hubs. This bodes well for incorporation of large proteins into the pSAGE particles, as well as for the self-repair of SAGE structures.

Protein-Hubs incorporate into pSAGEs

Next, we tested the assembly of modified SAGE particles by combining the HubA variants incorporating protein-fusions, parent HubA and K4-HubB in the ratio 1:9:10, in HEPES buffer at pH 7.2 and 20˚C. We refer to these as "5% pSAGE" as 5% of the CC-Tri3 motifs were fusion-proteins. We estimate that 5% incorporation equates to ≈75 copies of the fusion protein per 100 nm SAGE particle. The fusions used were either GFP-HubA or HubA-GFP to give GFP-SAGE and SAGE-GFP, respectively. As controls, two SAGE preparations were made, one without any fusion protein (parent-SAGE), and another with parent-SAGE mixed with free GFP at a concentration of 5% of CC-Tri3 (GFP-control).

As a straightforward test for SAGE formation, we prepared four 200 µl samples as above to total hub concentrations of 25 µM, and therefore 1.25 µM in the protein fusions. After incubation at 20˚C for one hour, the samples were centrifuged and the pellets were inspected by ambient and UV light, Figure 2a. Each sample containing both SAGE and GFP constructs gave green pellets. For each of these, we measured the fluorescence in the supernatant and compared it to 1.25 µM of the respective free GFP construct. For both the GFP-SAGE and SAGE-GFP preparations, ≥98% of the GFP signal was retained in the pellets. In contrast, for the GFP-control 28% of the fluorescence remained in the supernatant, Figure S5a. These data show that SAGE actively incorporate protein cargoes *via* fusion to the CC-Tri3 to make pSAGE particles. However, they also indicate that SAGE particles can either adhere to, or encapsulate, protein passively.

Figure 2. Characterization of assembled pSAGE particles by fluorescence, light scattering and microscopy. From top to bottom, direct visualization (**a**, labeled 'pellet'), dynamic light scattering (**b**, DLS), light microscopy (**c**, LM), scanning and transmission electron microscopy (**d**, SEM; **e**, negative-stain TEM) and atomic force microscopy (**f**, AFM) were used to visualize parent SAGE, the GFP-control, and the two pSAGE assemblies, GFP-SAGE and SAGE-GFP, data for which are presented from left to right. All samples were prepared with 5% of the

appropriate HubA—protein fusion to a final Hub concentration of 25 μM, except for DLS, which was conducted at 3 μM. For LM, pSAGE were assembled in the presence of 5% HubB-TAMRA to allow co-visualization of all SAGE particles through a red channel.

The four SAGE preparations were then compared by dynamic light scattering (DLS) at total hub concentration of 3 μM. Both control samples gave weak DLS autocorrelation functions, Figure S5b, and, consequently, broad and unreliable distributions of particle sizes, Figure 2b. We posit that the low signal intensities and poor autocorrelation functions for these control samples of unmodified SAGE particles arises because of the low density and, therefore, the poor scattering of peptide in the skin of the parent SAGE particles. By contrast, the GFP-SAGE and SAGE-GFP samples gave more-intense DLS signals with sharper peaks and tighter size distributions. From these data, the hydrodynamic radii for GFP-SAGE and SAGE-GFP particles were calculated to be 190 nm ± 15 nm and 138 nm ± 7 nm, respectively. The *N*-terminally decorated GFP-SAGE appeared larger than SAGE-GFP, which is consistent with our working hypothesis that the *N* termini of the CC-Tri3 units are presented predominantly on the outer surfaces of the assembled particles,[29] although the difference in sizes does not correspond simply to an additional layer of GFP molecules, which would be 4 – 6 nm deep.

The samples were imaged directly by wide-field fluorescence light microscopy (LM), Figure 2c. In each case, the SAGE particles were given a second color by incorporating 5% HubB-TAMRA (TAMRA = carboxytetramethylrhodamine) during assembly. Accordingly, the parent-SAGE preparation gave red puncta consistent with the presence of only TAMRA. Whereas, pSAGE assemblies were visible in both the red and green channels indicating incorporation of both TAMRA and GFP into SAGE. Indeed, the two signals were coincident demonstrating co-assembly of the different hubs into the same particles, Figures 2c and S6. Interestingly, the GFP-

control SAGE also gave green puncta albeit with a high level of green background. This suggests further that free GFP associates with the surfaces of SAGE particles.

Scanning and transmission electron microscopy (SEM and TEM) were also used for visualization, Figure 2d,e. For the former, samples were air dried on stubs and sputter coated with Au-Pd; while for the latter contrast was provided with 1% uranyl acetate as a negative stain. In both methods, low-magnification images of the samples revealed fields of SAGE particles, Figure S5d,e. At higher magnifications, separate particles were more abundant and more readily observed for the two pSAGE assemblies than in either control, Figure 2d. In negative-stain TEM, pSAGE particles were visible as defined particles, Figure 2e, and more clearly than we have observed previously for parent SAGE preparations. Samples were also inspected by Correlative Light Electron Microscopy (CLEM),[36] *i.e.*, imaged by LM followed by TEM. The superimposed images revealed overlaid particles and fluorescent puncta, Figure S5e, confirming that the particles visualized by EM do harbor GFP-fusion proteins.

To probe the non-specific interaction seen in the GFP-control, we exploited the increased contrast afforded to the SAGEs under TEM by associated protein. Assembly of SAGEs in the presence of bovine serum albumin (BSA, which has pI = 4.7, similar to GFP) and hen-egg lysozyme (HEWL, pI = 11.35) was investigated. As SAGE particles carry an overall positive charge at neutral pH, we assumed that BSA would bind to SAGEs but that HEWL would not. Consistent with this, parent-SAGE particles assembled and then treated with BSA were clearly visible by negative-stain TEM, but equivalent amounts of HEWL gave much less contrast, Figure S7. Thus, the passive association/encapsulation of proteins with/in SAGEs referred to above appears to correlate with charge, negatively charged proteins are more likely to interact

with the positively charged SAGE particles. However, this effect is weak compared with the active incorporation achieved when fusion proteins are included as components of the hubs.

We measured zeta potentials (ZP) to probe the overall effective charges on the particles through the above experiments (Figure S8). As anticipated from charges calculated using the polypeptide sequences, the ZP values for HubA and HubB were negative and positive, respectively; that for assembled SAGEs was positive; and those for the BSA and lysozyme additives were negative and positive, respectively. We found that when mixed with BSA the ZP of the SAGE particles shifted to the negative regime confirming a strong interaction between the particles and this protein (Figure S8b). However, the mixing experiment with HEWL gave a less-pronounced change in ZP indicating less interaction (Figure S8c).

Finally, atomic force microscopy (AFM) was used to visualize particles, and to give an alternative measure of particle dimensions, Figure 2f. Samples were deposited onto mica and the buffer wicked off to leave dried particles adhered to the substrate. The dimensions of the four SAGE preparations showed skewed distributions, Figure 3. The data raise a number of interesting points.

First, the diameters of all the particles were consistent with the DLS measurements for parent-SAGE and the protein-SAGE particles above. Second, with a 6-nm height cut-off to reduce noise (chosen as just less than twice the length of a CC-Tri3 module), no particles less than 50 nm in diameter were detected from any of the preparations, Figure 3a and Figure S9. This suggests a minimum size for mature SAGEs regardless of the decoration of the hubs. Turning to the height to diameter ratio measurements, all of the particles collapsed to some degree, Figure 3c. Although the heights of the particles varied both within each sample and between the four preparations, the thicknesses of the two protein-SAGE assemblies were larger than those of the

controls. Again, this is consistent with active decoration of SAGE particles adding bulk to the fabric of the SAGE particles. Interestingly, the height:diameter ratios were all ≈1:3, Figure 3c. To a first approximation, completely malleable, unilamellar objects might be expected to collapse down to the height of two times the lamellar thickness. That we do not see this suggests that the construction of the SAGE particles is more complicated than we had assumed previously, and hints at some internal structure, and/or that the decorated particles are stiffer or deform less uniformly than the non-decorated particles.

Figure 3. Analysis of single-particle measurements from atomic force microscopy (AFM). Box and Whisker plots for (**a**) the diameter, (**b**) the height and (**c**) the height to diameter ratio for parent-SAGE, GFP-control SAGE, GFP-SAGE and SAGE-GFP. (**d**) The total counts and median values for each of the above distributions.

In summary to this section, a battery of biophysical and microscopic techniques consistently shows that SAGE particles can be prepared actively incorporating folded and functional proteins. The resulting particles, which from here on we refer to as pSAGE, are easier to image than undecorated SAGE particles, and their sizes follow expectations from the working model for SAGE assembly. The improved ability to observe pSAGE particles, particularly by DLS and negative-stain TEM, is worth further comment: in our working model for SAGEs the hexagonal lattice of hubs is ≈ 3 nm thick and comprises only ≈40% peptide material. This gives low electron density for TEM imaging and hampers scattering needed for DLS. We posit that in the pSAGEs, pendant GFP molecules provide additional density to this peptide network, thereby improving its contrast and imaging. Related to this, the GFP-control are also more readily visualized, which suggests that parent SAGE particles can bind certain proteins non-specifically.

SAGE particles tolerate the incorporation of multiple copies of pendant proteins

Our working model for SAGE assembly is that GFP-SAGE particles present GFP on the outer surfaces of SAGEs and that SAGE-GFP encapsulates GFP molecules within the interiors of the particles. This follows from MD simulations of patches of the parent-SAGE particles.[29] To test the limits of SAGE load bearing, we extended these MD studies to 19-hexagon patches incorporating 5%, 15%, 33.3% and 50% of the GFP-HubA or HubA-GFP evenly spaced through the lattice, Figure 4a and Video S1-S4. Regardless of the fusion construct, the patches curved in the same direction as the unadorned patches,[29] resulting in the *N* termini of CC-Tri3 being presented on the convex (outer) faces. Thus, curvature appears to be an intrinsic property of the underlying peptide assembly. Calculations of the curvature of the patches suggested that increasing the amount of GFP-HubA could give a ≈4-fold increase in diameter of GFP-SAGE particles, Figure S10. Increasing loads of HubA-GFP, however, were anticipated to have a more dramatic effect, with the predicted diameters of the SAGE-GFP particles rising by almost 15-fold in analogous simulations. Thus, MD predicts that loading *C*-terminally appended protein into the pSAGE causes a greater expansion of the particles than introducing *N*-terminally fused protein.

We tested the loading capacity of the pSAGEs experimentally by visualizing particles with SEM. In addition to foregoing experiments at 0% and 5%, which yielded typical SAGE particles, samples were prepared at 15%, 25% and 35% of the fused hubs. Moving through these higher percentages, GFP-SAGE first aggregated together, then formed linked assemblies, and finally gave large, 1 μm – 3 μm diameter aggregates, Figure 4b. In comparison, the SAGE-GFP particles first increased in size, then aggregated, although to a lesser extent than the GFP-SAGE, and then formed what appeared to be sheet-like materials. Manual grain-sizing on SEM images of the preparations with 0%, 5% and 15% of the fusion hubs revealed that, with increasing load,

there was a greater increase in the diameter of the SAGE-GFP particles (134 nm – 199 nm) than for the GFP-SAGE particles (169 nm – 198 nm), Figure S11. These data agree at least quantitatively with the MD simulations.

Figure 4. Loading capacity of pSAGE particles. (a) Snapshots from 40 ns all-atom molecular-dynamic simulations of patches of initially flat hexagonally arrayed SAGE hubs incorporating regularly spaced GFP-HubA (left) or HubA-GFP (right) at 15% of the fusion protein. The simulations were conducted in aqueous buffer at 298 K. Videos of these simulations are available in the SI. **(b)** SEM images of pSAGE particles incorporating (from left to right) 0%, 5%, 15%, 25% and 35% of GFP-HubA (top) and HubA-GFP (bottom). Samples were prepared at 25 μM total hub concentration in 25 mM HEPES buffer, pH 7.2 at 20˚C and left to equilibrate for 60 minutes before preparation for imaging, which involved drying samples on mica and coating with ≈5 nm Au-Pd.

Based on the MD simulations and experiments, we propose that the introduced proteins provide additional steric bulk to the surfaces of SAGE particles, which, above certain thresholds, affects the curvature of the arrayed hubs and the appearance of the assembled particles dramatically. Moreover, because the intrinsic curvature has a defined direction, *N*- and *C*-terminally appended GFPs give different outcomes. Increasing the proportion of fusion protein in SAGE-GFP leads to reduced curvature in the MD simulations and then to flat arrays. This is manifest experimentally where high proportions of SAGE-GFP lead to large sheet-like materials rather than particles; effectively, increased *C*-terminal protein "pops" the SAGEs. The behavior of the GFP-SAGE is different, as particles appear to first aggregate and then form larger non-SAGE spherical particles. We rationalize this in two ways: first, the additional protein mass accentuates the intrinsic curvature of the underlying array; second, large amounts of externally facing GFPs leads to some aggregation of the particles.

Whilst the simulations provide only a semi-quantitative correlation to experiments—for example, the parent SAGE are predicted to have diameters of between 30 nm – 35 nm, based on

the 19 hexagon patches, rather than ≈100 nm observed—it is remarkable that they capture the experimentally observed differences between the GFP-SAGE and SAGE-GFP constructs.

Clearly, several factors influence SAGE formation, the sizes and properties of the resulting particles, and, indeed, whether spherical particles form at all. Nonetheless, the empirical data presented above show that *typical SAGE particles*, that is, spherical particles of ≈100 nm in diameter, still form with GFP fused to up to 15% of the CC-Tri3 peptides. This is extremely encouraging for future applications for confining and concentrating proteins within the SAGEs, as we estimate that 15% incorporation of protein fusion is an equivalent to a loading of ≈225 protein molecules per 100 nm particle. On this basis, we estimate an average effective concentration of low mM protein appended to SAGE particles. These results provide a strong basis for modeling more-complex SAGE-based systems in the future.

Simultaneous incorporation of multiple different proteins into pSAGEs

The co-incorporation of multiple protein functions into pSAGE particles will be key to realizing potential applications of these materials as delivery and vaccine platforms[37-38] and as bionanoreactors.[39] As a proof of concept for this, we tested the co-location of two different fluorescent proteins into SAGE particles. Through these studies, we found that *N*-terminally appended MBP improved the expression of many of our CC-Tri3-based constructs and the solubility of the resulting proteins.

An mCherry containing fusion was made to give MBP-mCh-CC-Tri3, which was used to generate the HubA variant, MBP-mCh-HubA. Two 3 μM SAGE preparations were then made with 5% MBP-mCh-HubA plus 5% of either GFP-HubA or HubA-GFP. Fluorescence

microscopy confirmed co-incorporation of the two fluorescent proteins into individual particles for both preparations, *i.e.,* signals from the red and green channels superimposed, Figure S12.

We tested for proximity of the two colored proteins in both dual-protein SAGE assemblies using Förster resonance energy transfer (FRET) between the GFP,[40] as the donor, and mCherry, as the acceptor, which has a Förster distance of ≈5 nm.[41] Control experiments with the two fusion HubA proteins in the presence of HubA, but the absence of K4-HubB gave minimal FRET. However, the FRET signal increased 3- or 6-fold for the SAGE preparations, Figure 5a and Figure S13. Next, we used fluorescent lifetime imaging microscopy (FLIM) to monitor the donor GFP directly. The fluorescent lifetimes of GFP in both GFP-SAGE and SAGE-GFP were reduced when the SAGE preparations included MBP-mCh-HubA, Figure 5b and Figure S14. This is best explained by the GFP transferring fluorescence energy to the mCherry with the consequence of reducing the fluorescence lifetime of the former. Thus, both the FRET and FLIM experiments are consistent with co-located and proximal GFP and mCherry fusion proteins. Interestingly, the larger changes in both experiments were observed for SAGEs made with HubA-GFP, which, by our working model, should place the GFP on the opposite side of the protein lattice from the mCherry. At this time, we cannot offer a simple explanation for this result. However, we note that detailed interpretations of FRET and FLIM data are complicated: we have shown that GFP adheres to SAGEs, and any such association between the pendant fluorescent proteins and the skin of the SAGEs may impose some orientation on the fluorescent proteins; in turn, this may affect the FRET and FLIM measurements. However, this is purely speculative at the moment.

Figure 5. Exploring the requirements for developing pSAGEs into nano-reactors. (a) Förster resonance energy transfer (FRET) intensity measurements for samples containing 5%

MBP-mCh-HubA with either 5% GFP-HubA or HubA-GFP without K4-HubB, not forming SAGE particles (dark grey) and with K4-HubB, forming SAGE particles (light grey). (**b)** Fluorescence lifetime imaging microscopy (FLIM) measurements for 5% GFP-SAGE and SAGE-GFP (dark grey) and for preparations also containing 5% MBP-mCh-HubA (light grey). **(c)** Bioluminescent emission at 472 nm from the catalysis of coelenterazine by *Renilla* luciferase for 5% MBP-Luc-HubA and MBP-HubA-Luc in the absence of K4-HubB, which does not form SAGE particles (dark grey); and with K4-HubB, which forms SAGE particles (light grey).

Finally, we incorporated the luciferase enzyme (Luc) into SAGE particles, Figure 5c. Two fusion proteins were made, MBP-Luc-CC-Tri3 and MBP-CC-Tri3-Luc, we used these to make HubA variants, and then assembled MBP-Luc-SAGE and MBP-SAGE-Luc, respectively, Figure S1. The luciferase activities of these constructs were compared with that for the free fusion proteins by adding coelenterazine and monitoring at 472 nm over 200 s. The activities of both SAGE assemblies were reduced compared with the free enzymes: MBP-Luc-SAGE retained 75% ±4% activity, and MBP-SAGE-Luc retained 64% ±7% activity. Thus, and critically, both luciferase—SAGE constructs are active. Whilst it is tempting to interpret the small difference between the activities of two constructs in terms of substrate access to Luc presented on the outside and the inside of the SAGE particles, respectively, we note that these activities are within the experimental errors for these measurements.

Together, the fluorescence microscopy, FRET, FLIM, and luciferase data provide strong evidence that multiple copies of different proteins can be successfully co-located into pSAGEs. In addition, enzyme immobilization by SAGE particles only mildly impairs catalytic activity, both illustrating that active proteins can be incorporated to SAGEs and that these constructs still permit substrate access to the enzymes. This is encouraging for future work that aims to exploit the SAGEs as delivery and encapsulation vehicles for bioactive proteins.

Conclusion

Here, we have shown that proteins can be actively incorporated into self-assembled cage-like particles, SAGEs, to form protein-SAGE constructs. This is done by making fusion proteins comprising the target protein and one of the three peptide modules of the SAGE system. These fusions are used to make two types of complementary hub that are the components for SAGE assembly. Biophysical measurements show that fusions assemble into hubs, and light and electron microscopy confirm the incorporation of these into larger, 100 nm – 200 nm particles. Incorporation of the target proteins is near complete, and the general morphology of the resulting pSAGE particles is maintained provided that the proportion of fusion protein is kept to ≈15%. At this level of incorporation, and for 100 nm diameter pSAGEs, ≈225 protein-fusions are incorporated within a volume of ≈500 zeptolitres, translating to a local concentration of ≈1 mM protein. Proteins that have been incorporated into pSAGEs thus far include: fluorescent proteins, solubilizing globular proteins, and an enzyme. They can be fused to either the *N* or *C* terminus of the peptide building block giving some control over whether the protein is displayed mainly on the outside of particles, or encapsulated within them, respectively. Finally, multiple different protein fusions can be incorporated into the SAGE particles simultaneously.

The advantages of this system are: (1) the recombinant production of the pSAGE components makes a wide range of target proteins accessible; (2) although some manipulation of the protein fusion is needed, *i.e.* a disulfide linkage to complete the hubs, this is minimal and uses standard chemistry in biological buffers; (3) the SAGE system offers control over the orientation and stoichiometry of the target proteins; (4) the modularity of the SAGE system allows target proteins to be incorporated within pSAGE assemblies rapidly and with ease; (5) following these features, fusion proteins are integrated into the fabric of the pSAGE rather than being passively associated or encapsulated; (6) as the SAGE particle surface is only ≈40% peptide, and has ≈6

nm pores, good mass transfer of small-molecule analytes and substrates should be possible across this skin; and (7) because the pSAGEs are ≈100 nm in size they can be employed in suspension, or as a gel/solid after mild centrifugation. For these reasons, we believe that the pSAGE will find use as nanoscale materials for applications in biotechnology, nanotechnology and synthetic biology; for instance, the encapsulation of enzymes and enzyme cascades.

On this last point, benefits of enzyme immobilization include: facilitating the separation of products from the active enzymes; and increasing enzyme shelf life by protecting them against proteolysis, thermal and chemical denaturation.[42-44] There are several methods for the industrial immobilization of enzymes, such as: covalent linkage[44] or adsorption[45] (generally ionic) to a support; entrapment within a matrix;[46] and direct enzyme crosslinking.[47] There are disadvantages of these current methods, however, which include: the leakage of the protein from supports; loss of enzyme activity due to protein unfolding or misfolding; and restricted protein mobility and/or access of substrates due to inappropriate cross-linking. Soluble, nanoscale supports, such as the pSAGE, which immobilize enzymes through understood protein-protein interaction domains, potentially overcome some of these shortcomings. They allow for high densities of enzymes to be achieved actively and prescriptively, and without compromising enzyme structure or activity. In turn, these could help increase substrate channeling between multiple enzymes in a cascade.[48,49] Furthermore, encapsulation of enzymes within structures that isolate them from bulk solvent can impart enhanced thermostability and increased tolerance to proteases.[50] Given the above development of pSAGEs, and this potential for active encapsulation or presentation of enzymes, our next step will be to incorporate multiple enzymes of a catalytic pathway into the same SAGE particle to produce enzyme nano-reactors, or eSAGEs.

Materials and methods

Parent-SAGE component synthesis

The synthesis of peptides by solid-phase peptide synthesis (SPPS) and the formation of hub molecules, through an asymmetric disulfide bond, followed the protocol described by Fletcher *et al*.[29] with the following exceptions. (1) The CC-Tri3 resin was split after synthesis, half of which was modified with an additional KKKKGG on the *N* terminus, compared to the standard CC-Tri3 molecules, to give K4-CC-Tri3 . (2) Thiol activation of the cysteine by 2,2'-Dipyridyldisulfide (DPDS) was performed on CC-DiA and CC-DiB, to give CC-DiA(SPy) and CC-DiB(SPy), as opposed to CC-Tri3(Spy). (3) HubA was formed by combining CC-Tri3 and CC-DiA(SPy), K4-HubB was formed by combining K4-CC-Tri3 and CC-DiB(SPy), all associated procedures follow techniques outlined by Fletcher *et al*.,[29]. Successful synthesis and characterization of K4-HubB, previously unreported, can be found in the supporting information (Figure S2).

Molecular Biology

Protein sequence information is available in the supporting information (Figures S15-22). The design of the GFP containing fusion proteins, plasmids and analysis of sequencing data were conducted on GeneDesigner[51], Bioedit[52], and Benchling[53]. These *in silico* sequences were further optimized, synthesized, expressed and purified by GenScript to give GFP-CC-Tri3, CC-Tri3-GFP, MBP-GFP-CC-Tri3 and MBP-CC-Tri3-GFP. The mCherry and luciferase sequences were ligated into these original plasmids *via* an EcoRI-KpnI digest (GFP-CC-Tri3 and MBP-GFP-CC-Tri3) or a SalI-NheI digest (CC-Tri3-GFP and MBP-CC-Tri3-GFP). All restriction enzymes were acquired from New England Biolabs, as HF versions if available. Vectors

containing the mCherry gene and Renilla luciferase were obtained from Professor Paul Verkade, University of Bristol and GSK respectively. All vectors were initially transformed into *E. coli* XL10 Gold (Stratagene, 200314) from which working plasmid stocks were derived, these were then used to transform *E. coli* BL21 (DE3) cells (New England Biolabs, C25271).

Protein expression and purification

E. coli BL21(DE3) stocks of a chosen plasmid were used to inoculate 400 ml of LB media in 2 L Erlenmeyer flasks with 50 µg/mL of Kanamycin, which were incubated (Thermo Scientific, MaxQ 4000) at 37°C and shaken at 200 rpm. Growth was monitored until 0.6-0.8 absorbance at OD_{600} nm was reached, at which point gene expression was induced *via* the addition of 500 µM IPTG (final concentration). After induction, cells were incubated at either 18°C, 25°C or 37°C, 200 rpm, overnight (≈16 hours).

Induced cultures were centrifuged for 10 minutes at 10'000 x g (Thermo Scientific, Sorvall Lynx 4000) and the cell pellet was transferred into 50 ml centrifuge tubes, 25 ml of lysate buffer (Phosphate buffered saline (PBS), pH 7.4, 20 mM imidazole) was used to resuspend the cell pellet, which was then sonicated (BioLogics, model 3000) for 10 minutes on ice. The cell lysate was then centrifuged for 30 minutes at 29'000 x g and the lysate was clarified through a 0.45 µm syringe filter. The lysate was applied to a 5 ml HisTrap HP [GE Life Sciences, 17-5248-01) *via* an ÄKTAprime (GE Life Sciences, ÄKTAprime plus). The column was washed with 25 ml wash buffer (PBS, pH 7.4, 50 mM imidazole) before it was eluted with 25 ml his-tag elution buffer (PBS, pH 7.4, 300 mM imidazole), and the fractions collected. Proteins containing MBP were pooled and applied, *via* the ÄKTAprime, to an MBPtrap HP (GE Life Sciences, 28-9187-79) column, which was washed in 25 ml PBS, pH7.4 and eluted with MBP elution buffer (PBS, pH7.4, 10 mM maltose). Eluted fractions at 300 mM imidazole or 10 mM

maltose were pooled and concentrated to 2-3 ml before size exclusion chromatography (SEC) was performed by way of the ÄKTAprime on a Superdex 200 pg HiLoad 16/600 (GE Life Sciences, 28-9893-35) in PBS, pH7.4. Fractions which eluted at a volume appropriate for the molecular weight anticipated were visualized by SDS PAGE (UVP, BioDockit) and pure samples were pooled, if appropriate these were flash frozen in liquid nitrogen and stored at -80°C.

Modification of fusion proteins by activated CC-DiA

Purified proteins in PBS were mixed in a 1:2 molar ratio with CC-DiA(SPy) at a range of concentrations dependent on the protein in question. The solutions were agitated for at least 2 hours at 20°C or overnight at 4°C, after which excess CC-DiA(SPy) was removed *via* an ÄKTAprime with either SEC (Superdex 200 pg HiLoad 16/600 (GE Life Sciences, 28-9893-35)) or with a salt exchange column (HiTrap Desalting (GE Life Sciences, 29-0486-84)) during which the modified protein was exchanged into 25 mM HEPES at pH 7.2. Mass spectrometry was conducted on both unmodified and modified proteins to determine the success of the modification *via* MALDI-TOF (Bruker, UltraFlex). Samples were prepared after isolation of the protein from the buffer *via* the use of C18 reverse phase chromatography (Millipore, ZipTips C18 P10) and eluted with 10 μM 0.1% trifluoroacetic acid in 50% acetonitrile. These samples were mixed 2:1 or 4:1 with 2,5-dihydroxybenzoic acid (DHB). See supplementary information for protein characterization, Figures S15-S22. BSA (sigma, MW 66430) was used as a standard for MALDI-TOF at these higher molecular weights. An average of 66229.97 Da ± 41 Da was recorded over six measurements, giving an accuracy of 0.3%.

K4-CC-Tri3 and K4-HubB characterization

For HPLC, crude K4-CC-Tri3 peptides were purified to homogeneity by reverse-phase high-performance liquid chromatography (RP-HPLC) in a semi-preparative (Phenomenex Kenetic (5 μm, 100 Å, 10 mm ID x 150 mm L) C18 reverse phase column, 3 mL per min flow rate) manner employing 0.1% TFA in H_2O (A) and 0.1% TFA in MeCN (B) as eluents. A linear gradient of 20% to 80% B was applied over 40 min. Collected fractions were analyzed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and automated analytical RP-HPLC (with a Phenomenex Kenetic (5 μm, 100 Å, 4.6 mm ID x 100 mm L) C18 reverse phase column, 1 mL per min flow rate (otherwise as above)) before fractions found to contain solely the desired product were pooled and lyophilized, Figure S2a and b.

For circular dichroism (CD) spectroscopy, K4-CC-Tri3 was analyzed in phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM phosphate buffer, pH 7.4) in 1 mm path length quartz cuvettes using a JASCO J-810 or J-815 spectropolarimeter fitted with a Peltier temperature controller. K4-CC-Tri3 was analyzed with a 5-fold excess of TCEP (a potent reducing agent) to prevent disulfide bond formation at high temperatures.[54] Thermal denaturation experiments were performed by increasing the temperature from 5°C to 90°C at a linear rate of 40°C per hour with full spectra recorded at 5°C intervals and the circular dichroism at 222 nm recorded at 1°C intervals. All raw data were normalized for concentration, path length, and number of amide bonds present. Melting temperatures (T_M) were determined from the point of inflection of a thermal denaturation curve, Figure S2e and f

Sedimentation equilibrium (SE) analytical ultracentrifugation (AUC) experiments were performed at 20°C in a Beckman Optima XL-A analytical ultracentrifuge employing an An-50 Ti rotor with Epon 6 channel centrepieces and quartz windows. K4-CC-Tri3 was analyzed in PBS at 325.5 μM concentration with a 5-fold excess of TCEP to prevent disulfide bond

formation over the experiment. Reference channels contained PBS. Samples were centrifuged at speeds in the range of 22 - 42 krpm. Collected data was fitted to a single, ideal species model using Ultrascan II and 95% confidence limits were calculated by *Monte Carlo* analysis of the obtained fits, Figure S2d.

Functionalization of HubB with TAMRA

To introduce carboxylic-acid functionalized carboxytetramethylrhodamine (TAMRA) to the CC-Tri3 *C* terminus an orthogonally-protected lysine residue (Fmoc-Lys(alloc)- OH) was employed during SPPS. To selectively deprotect the alloc group the resin was thoroughly washed with deoxygenated dichloromethane (DCM) and incubated with deprotection mix (1 eq. $Pd(PPh_3)_4$, 40 eq. phenylsilane per 1 eq. peptide resin, 10 mL degassed DCM, 30 mins, RT). The resin was washed with 3 x 20 ml DCM, 3 x 20 ml with DMF and 3 x 20 ml with deoxygenated DCM before incubated with fresh deprotection mix (30 mins, RT). The resin was washed 3 x 20 ml with DCM, 3 x 20 ml with DMF, 2 x 20 ml with dioxane:H2O (9:1, v/v), 20 ml with MeOH and 3 x 20 ml with DMF. Dye molecules were coupled through amine bond formation chemistry (1 eq. peptide, 1.5 eq. TAMRA, 1.35 eq. hydroxybenzotriazole (HOBt), 1.5 eq. N,N'-diisopropylcarbodiimide (DIC) in DMF, RT, 12 h). This CC-Tri3-TAMRA molecule was converted to HubB-TAMRA *via* formation of a disulfide bond, as described above, Figure S2c.

Characterization of protein-hub oligomerization state and trimer exchange.

Spectra from circular dichroism (CD) were obtained from 5-90°C with a JASCO J-810 spectropolarimeter. 50 mM of protein sample in 50 mM potassium phosphate, pH 7.4 (16.04 mM dibasic K2HPO4, 3.96 mM monobasic KH2PO4) was analysed in a 0.1 cm path length quartz cuvette. Absorbance at 222 nm was monitored every 1 °C, and a CD spectra between 260–190 nm was recorded every 5 °C.

Size exclusion chromatography was conducted on an ÄKTAprime with Superdex 200 pg HiLoad 16/600 (GE Life Sciences, 28-9893-35) in 25 μM HEPES pH7.2. SEC of 500 μl of 25 μM HubA-GFP was performed followed by SEC of 500 μl of 250 μM HubA. SEC was then performed on a solution of 500 μl containing both 25 μM HubA-GFP and 250 μM HubA that was injected immediately upon mixing. Finally SEC was conducted on a solution of 500 μl containing both 25 μM HubA-GFP and 250 μM HubA which was left to incubate at 20°C for 2 hours. All components were prepared in 25 μM HEPES pH 7.2

Analytical ultracentrifugation (AUC) sedimentation velocity experiments were performed at 20 °C on a Beckman Optima XL-A ultracentrifuge with an An-50 (Ti) rotor (Beckman-Coulter). Experiments were performed in a sedimentation velocity cell with a 2-channel charcoal centre piece and quartz windows. CC-Tri3-GFP (50 μM, in 25 mM HEPES) and 25 mM HEPES (420 μ L) were placed in the sample and reference channel respectively. Cells were centrifuged at 35 krpm and absorbance scans taken at 5-minute intervals for 120 scans. The buffer density and partial specific volume of CC-Tri3-GFP were calculated using SEDNTERP.[55] The baseline, meniscus, frictional coefficient (f/f0) and systematic time-invariant and radial-invariant noise were fitted to a continuous c(s) distribution using SEDFIT,[56] at 95% confidence level.

General SAGE and pSAGE assembly

SAGE and pSAGE particles are composed of two hub solutions, HubA and HubB. Each of these hub solutions can be doped with the modified hubs (either protein or TAMRA *etc*) before the solutions are then combined to form SAGE or pSAGE particles. When SAGEs are prepared with modified hubs, the doped HubA and HubB solutions are incubated for 20 minutes to equilibrate prior to mixing. Unless otherwise stated, all SAGEs are doped with 10% of the modified hubs compared to the individual hub solutions; hence, when HubA and HubB solutions

are mixed, 5% of the total CC-Tri3 motif in solution presents the said modification, all SAGE particles were made in 25 μM HEPES at pH 7.2. For example, to produce 200 μl of 25 μM, 5% SAGE-GFP: (1) mix 10 μl of 25 μM HubA-GFP with 90 μl of 25 μM HubA, allow 20 minutes incubation at 20°C, this is the HubA solution: (2) mix 100 μl of HubA solution with 100 μl of 25 μM HubB, HubB does not require the 20 minute incubation as it has a single constituent: (3) incubate the SAGE particles at 20°C for 1 hour to complete SAGE particle formation: (4) this gives a final concentration of 1.25 μM of the fusion protein. Given that the GFP-control protein does not partake in the SAGE lattice structure, SAGE preparations with the GFP-control protein are 1:1 HubA to HubB with the GFP-control protein added, to give the same final relative concentration of protein as the pSAGE mixture.

pSAGE characterization

Centrifugal pelleting experiments to determine the completeness of protein incorporation were performed on 200 μl of SAGE particles at 25 μM and on 200 μl of 5% pSAGE particles at 25 μM, doped with either GFP-HubA, HubA-GFP, MBP-GFP-HubA or MBP-HubA-GFP (as described above), additionally a further SAGE sample was prepared with 1.25 μM GFP-control. In addition, another set of preparations were made replacing HubB with buffer as negative controls. After the 1-hour incubation, all tubes were centrifuged at 6000 x g for 6 minutes (Thermo Scientific, Heraeus Pico 17). Images were taken immediately after centrifugation under UV radiation in both a darkened room and in ambient light. Supernatants were then extracted from the tubes and the fluorescence was measured using a Spectrofluorometer (Jasco, FP-6500)

Dynamic light scattering (DLS) experiments, to determine the hydrodynamic radii of pSAGE particles, were performed on 200 μl of SAGE particles at 3 μM and on 200 μl of 5% pSAGE particles at 3 μM, doped with either GFP-HubA or HubA-GFP (as described above),

additionally a further SAGE sample was prepared with 150 nM GFP-control. After 1 hour at 20°C, the samples were analyzed by DLS (Malvern, Zetasizer Nanoseries), for 10 replicates of 3 measurements for 12 seconds each, the data was processed by the Malvern software *via* 'protein analysis' (non-negative least squares analysis followed by L-curve).

Light microscopy (LM) observations were conducted to visualize GFP bound to pSAGE particles and confirm that particles consisted of both HubA and HubB. Thus all SAGE formulations contained 5% total HubB-TAMRA in addition to the following: 25 µl of SAGE particles at 25 µM and 25 µl of 5% pSAGE particles at 25 µM, doped with either GFP-HubA or HubA-GFP (as described above) were prepared, additionally a further SAGE sample was prepared with 1.25 µM GFP-control. Each preparation was incubated for 1 hour at 20°C for the SAGE particles to complete formation. 5 µl of the SAGE sample was added to glass-bottomed microscopy dishes (CELLview, cell culture dish, PS, 35/10 mm, glass bottom), spread out over one half of the dish and left to dry. The fluorescence of SAGE particles was then imaged by LM (Leica TCS SP8 attached to a Leica DMi8 inverted microscope). Images were evaluated using FIJI[57] and ImageJ[58]. All images were taken using the same parameters on the microscope, and thus, the images in Figure S6 are directly comparable.

Scanning electron microscopy (SEM) observations, to visualize the polydispersity of SAGE populations, were performed on preparations of 25µl SAGE particles at 25 µM and on 25 µl of 5% pSAGE particles at 25 µM, doped with either GFP-HubA or HubA-GFP, as described above. Additionally, a further SAGE sample was prepared with 1.25 µM GFP-control. After 1 hours incubation at 20°C, 5 µl of sample was applied to freshly split mica attached to an aluminium SEM stub and allowed to dry. A ≈5 nm coating of Au-Pd was applied to the samples by a

sputter coater (Emtech, 575X). Samples were visualized with a SEM (FEI, Quanta 200 FEG-SEM) with a voltage of between 5-15 keV. Images were evaluated using FIJI and ImageJ.

Transmission electron microscopy (TEM) observations, to visualize high-resolution SAGE structures and for use in CLEM analysis, were performed on the same preparations that were used in both the SEM and LM experiments described above. The solutions used in the SEM experiment were applied to TEM grids as described in the LM experiments. LM experiments were conducted before uranyl acetate staining. Given preparation of the TEM grid, and LM if required, 5 µl of 1% uranyl acetate was added to the grids and immediate wicked off, grids were then left for 30 minutes to dry. These were then imaged with a TEM (FEI, 120kV BioTwinSpirit). Images were evaluated using FIJI and ImageJ, CLEM images were correlated with the TurboReg plugin within FIJI.

Atomic force microscopy (AFM) observations, to determine particle diameters and heights, were performed on the same preparations that were used in the TEM, SEM and LM experiments above. 10 µl of solution was pipetted onto freshly cleaved muscovite mica, the sample was left for 5 min before being washed with 3mL of water and dried under a flow of nitrogen. Images were obtained using a Bruker Multimode AFM with Nanoscope V controller in tapping mode. The cantilevers used were Bruker Scanasyst-air- HR (resonance 130 kHz, spring constant 0.4 N/m, nominal tip radius 2 nm. All measurements were taken under ambient conditions. A 50 μm^2 image with a resolution of 5120 x 5120 pixels was recorded for each sample. The images were analyzed using Nanoscope analysis software.

Further TEM was conducted on 25 µM SAGE particles assembled in the presence of 1.25 µM of either lysozyme (Sigma-aldrich, 62970-5G-F) or bovine serum albumin (BSA, Sigma-aldrich,

B4287-5G). These samples were prepared and visualized in an analogous manner to the GFP-control SAGE preparations, but replacing GFP-control for lysozyme or BSA.

Zeta potentials (ZP) experiments, to determine the ZP at the slipping plane of pSAGE particles, were performed using 900 µl samples of 3 µM SAGE peptide components, in which 5% of HubA was replaced with either GFP-HubA or HubA-GFP (0.15 µM fusion-protein). In addition, SAGE samples were prepared with 0.15 µM GFP-control, BSA or HEWL. After 1 hour assembly time at 20 °C, the samples were placed in disposable folded capillary zeta cells (Malvern, DTS1070) and analyzed (Malvern, Zetasizer Nanoseries) at 25 °C with 12-100 measurements and repeated 3 times. ZP values were also measured similarly for the separate components at 3 µM concentrations.

Further SEM was conducted using increasing percentages of either GFP-HubA or HubA-GFP instead of parent-HubA, which were prepared to a final protein percentage of 5%, 15%, 25% and 35%. Otherwise these samples were assembled and treated identically to previous SEM experiments.

Molecular Dynamics

System setup: The 19 hexamer patch previously described[29] was separated into numbered files containing individual HubA trimer helices, HubB trimer helices, acidic and basic helices. The conjugated protein and linker was built on to a standard hub helix in InsightII. The array of hub trimer helices for replacement by the conjugated helix was selected to provide an even spread across the surface of the patch. InsightII was used to overlay the conjugated helix onto each of the standard helices selected to be "doped" in to the patch. A fortran program was written to reassemble the remaining hubs and helices to approximate the original 19 hexamer patch. Another in-house fortran program was used to fix minor discrepancies in disulfide bond lengths

by moving the associated cysteine SG atoms towards each other giving an S-S distance of 2.0 Å . This enabled the patch to retain the required 144 disulfides during the pdb2gmx process. Hydrogen atoms were added consistent with pH 7 and parameterized with the Amber-99SB-ildn forcefield. Each complex was surrounded by a box 4 nm larger than the polypeptide in each dimension, and filled with TIP3P water. Random water molecules were replaced by sodium and chloride ions to give a neutral (uncharged overall) box and an ionic strength of 0.15 M. Each box contained between 5 – 7 million atoms depending on the conjugated protein and "doping" levels. Each patch was subjected to 5000 steps of energy minimization prior to the molecular dynamics simulations. This procedure was sufficient for doping the patches with 10 to 30% conjugate hub trimer, but beyond that clashes made it energetically impossible to minimize the system. For these systems a tapered scaling was applied between the trimer helix (which remained unchanged) and the linker and conjugate such that the conjugate protein size was reduced in x and y by up to 50%. The system was restored to normal size over an extended minimization run of 100000 steps. This gradual re-expansion process allowed the individual proteins to occupy available space without encountering major energetically unfavorable clashes.

Simulation details: All simulations were performed as NPT ensembles at 298 K using periodic boundary conditions. Short range electrostatic and van der Waals' interactions were truncated at 1.4 nm while long range electrostatics were treated with the particle-mesh Ewald's method and a long range dispersion correction applied. Pressure was controlled by the Berendsen barostat and temperature by the V-rescale thermostat. The simulations were integrated with a leap-frog algorithm over a 2 fs time step, constraining bond vibrations with the P-LINCS method. Structures were saved every 0.1 ns for analysis and each run over 20 ns. Simulation data were

accumulated on the UK supercomputer Archer and the Bristol BrisSynBio supercomputer Bluegem.

Analysis: Curvature of the patches. The SG atoms of the cysteine residues were fitted to a sphere while allowing the radius and center to move in space. The procedure was implemented in python and ten repeats of the fitting procedure from random starting values of radius and center position was sufficient to identify the best fit.

Software: The GROMACS-4.6.7 suite of software was used to set up and perform the molecular dynamics simulations. Molecular graphics manipulations and visualizations were performed using InsightII, VMD-1.9.1 and Chimera-1.10.2. The accompanying videos created with PyMol (1.7.4.0 Open-Source), ffmpeg (version 2.5.4) and Handbrake (version 1.0.2)

FRET and FLIM

Förster resonance energy transfer (FRET) experiments, to understand the proximity between pendant proteins were performed on 100 µl of pSAGE particles at 3 µM, doped with 5% of either GFP-HubA or HubA-GFP and either with or without 5% MBP-mCh-HubA. Additionally, samples were made which lacked the K4-HubB component and therefore did not form SAGE particles. After 1 hours incubation at 20°C, the samples were excited at the GFP excitation maximum (495 nm) and emission was measured at the mCh emission maximum (603 nm) using a Spectrofluorometer (Jasco, FP-6500).

Fluorescence Lifetime Imaging (FLIM) was used to measure changes in Förster Resonance Energy Transfer (FRET) on 50 µl of 3 µM pSAGE samples. pSAGE particles contained 5% of either GFP-HubA or HubA-GFP both with, and without 5% MBP-mCh-HubA . Fluorescence lifetime images were acquired on a Leica TCS SP8 system attached to a Leica DMi8 inverted microscope (Leica Microsystems). Excitation was provided by a white light laser with a

repetition rate of 20 MHz and an acousto-optical beam splitter (AOBS) selected an excitation wavelength of 488 nm. Images were acquired using a 20x 0.75 NA air immersion objective. Fluorescence of the GFP was detected using a hybrid detector operating in photon counting mode over an emission range of 500 – 530 nm. A notch filter centered on 488 nm minimized any laser scatter into the detector. Time resolved data was acquired through use of a PicoHarp 300 TCSPC module (PicoQuant) controlled through SymPhoTime64 software (PicoQuant). FLIM Images were acquired with 256 x 256 pixels and 4096 time bins. Fitting of FLIM images was performed with the FLIMfit software tool developed at Imperial College London[59]. Temporal binning of the fluorescence decays was performed prior to fitting resulting in 256 time bins per decay. Global Analysis fitting of the images was then performed with a double exponential model on all pixels above an intensity threshold of 100 photons allowing spatial variations in the intensity weighted mean fluorescence lifetime to be visualized.

Luciferase assay

Luciferase bioluminescence experiments were performed to assess the effect SAGE immobilization has on enzyme kinetics. Preparations of 400 µl of pSAGE particles at 2 µM, doped with 5% of either MBP-Luc-HubA or MBP-HubA-Luc were assembled, additional samples which lacked the K4-HubB component and therefore did not form SAGE particles were also made. SAGE particles were incubated for 1 hour at 20°C. Assay conditions contained a final concentration of 1 µM pSAGE (thus 50 nM luciferase), 4 µM coelenterazine, 25 µM HEPES, pH 7.2 with a volume of 200 µl. Upon the addition of coelenterazine, luminescence from the product, coelenteramide, was measured at 472 nm using a spectrofluorometer (Jasco, FP-6500) for 200 seconds. Measurements were taken in triplicate and total cumulative emission recorded.

ASSOCIATED CONTENT

Supporting Information.

This material is available free of charge *via* the Internet at http://pubs.acs.org.

JFR-pSAGE-SI: Supporting experimental figures, sequence information and characterization data for protein constructs (PDF).

Video JFR-pSAGE-SI_V1: Video of GFP-SAGE_5% MD trajectory (AVI)

Video JFR-pSAGE-SI_V2: Video of GFP-SAGE_15% MD trajectory (AVI)

Video JFR-pSAGE-SI_V3: Video of SAGE-GFP_5% MD trajectory (AVI)

Video JFR-pSAGE-SI_V4: Video of SAGE-GFP_15% MD trajectory (AVI)

AUTHOR INFORMATION

Corresponding Authors

Derek N. Woolfson, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK. D.N.Woolfson@bristol.ac.uk

David Tew, GSK, Gunnels Wood Rd, Stevenage, SG21 2NY, UK. david.tew@gsk.com

Author Contributions

J.F.R., J.M.F., A.B., C.M.E., D.T. and D.N.W. conceived the project and designed the experiments. D.S. and R.B.S. performed the molecular dynamics simulations. D.A. conducted the FLIM microscopy. H.E.V.B. conducted the AFM experiments and analysis. J.L.B. made HubB-TAMRA. W.M.D. performed analytical ultracentrifugation on the CC-Tri3-GFP. J.F.R.

performed all other experimental work. J.F.R. and D.N.W. wrote the paper. All authors have read and contributed to the preparation of the manuscript.

Funding Sources

J. Ross, D. Tew, A. Bridges, C. Edge and D. Woolfson are supported by GSK. J. Fletcher and D. Woolfson are supported by the BBSRC grant (BB/L010518/1); H. Bray and D. Woolfson are supported by the ERC (340764); J. Beesley is supported by the BBSRC South West Doctoral Training Partnership; L. Hodgson is supported by a BBSRC grant to P. Verkade and D. Woolfson (BBM002969/1); D. Alibhai, J. Mantel and P. Verkade are members of the Wolfson Bioimaging Facility; D. Shoemark, R. Sessions and D. Woolfson are members of BrisSynBio, a BBSRC/EPSRC-funded Synthetic Biology Research Centre (BB/L01386X/1). D. Woolfson is a Royal Society Wolfson Research Merit Award holder (WM140008). We would also like to thank: the University of Bristol School of Chemistry Mass Spectrometry Facility for access to the EPSRC-funded Bruker Ultraflex MALDI TOF/TOF instrument (EP/K03927X/1); the EPSRC for awarding HECBiosim and an Archer Leadership Award for providing compute time on the UK supercomputer Archer; BrisSynBio for providing BlueGem HPC and the Bristol ACRC for hosting it; BrisSynBio for the FLIM and acknowledge the MRC; and the Wolfson Foundation for funding the University of Bristol's Bio-imaging Facility.

ACKNOWLEDGMENT

We thank: R. Harniman for help with the atomic force microscopy; N. Linden for the spherical curvature fitting program; and members of the Wolfson Imaging Facility and the Woolfson group for helpful discussions.

REFERENCES

1. Chessher, A.; Breitling, R.; Takano, E., Bacterial Microcompartments: Biomaterials for Synthetic Biology-Based Compartmentalization Strategies. *ACS Biomater. Sci. Eng.* **2015,** *1*, 345-351.

2. Cornejo, E.; Abreu, N.; Komeili, A., Compartmentalization and Organelle Formation in Bacteria. *Curr. Opin. Cell Biol.* **2014,** *26*, 132-138.

3. Diekmann, Y.; Pereira-Leal, J. B., Evolution of Intracellular Compartmentalization. *Biochem. J.* **2013,** *449*, 319-331.

4. Harold, F. M., Molecules into Cells: Specifying Spatial Architecture. *Microbiol. Mol. Biol. Rev.* **2005,** *69*, 544-564.

5. Küchler, A.; Yoshimoto, M.; Luginbühl, S.; Mavelli, F.; Walde, P., Enzymatic Reactions in Confined Environments. *Nat. Nanotechnol.* **2016,** *11*, 409-420.

6. Parsons, J. B.; Frank, S.; Bhella, D.; Liang, M.; Prentice, M. B.; Mulvihill, D. P.; Warren, M. J., Synthesis of Empty Bacterial Microcompartments, Directed Organelle Protein Incorporation, and Evidence of Filament-Associated Organelle Movement. *Mol. Cell* **2010,** *38*, 305-315.

7. Fan, C.; Cheng, S.; Liu, Y.; Escobar, C. M.; Crowley, C. S.; Jefferson, R. E.; Yeates, T. O.; Bobik, T. A., Short N-Terminal Sequences Package Proteins into Bacterial Microcompartments. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107*, 7509-7514.

8. Lawrence, A. D.; Frank, S.; Newnham, S.; Lee, M. J.; Brown, I. R.; Xue, W.-F.; Rowe, M. L.; Mulvihill, D. P.; Prentice, M. B.; Howard, M. J., Solution Structure of a Bacterial Microcompartment Targeting Peptide and Its Application in the Construction of an Ethanol Bioreactor. *ACS Synth. Biol.* **2014,** *3*, 454-465.

9. Zschoche, R.; Hilvert, D., Diffusion-Limited Cargo Loading of an Engineered Protein Container. *J. Am. Chem. Soc.* **2015,** *137*, 16121-16132.

10. Azuma, Y.; Zschoche, R.; Tinzl, M.; Hilvert, D., Quantitative Packaging of Active Enzymes into a Protein Cage. *Angew. Chem. Int. Ed.* **2016,** *55*, 1531-1534.

11. Putri, R. M.; Cornelissen, J. J.; Koay, M. S., Self-Assembled Cage-Like Protein Structures. *Chemphyschem* **2015,** *16*, 911-918.

12. Bode, S. A.; Minten, I. J.; Nolte, R. J. M.; Cornelissen, J. J. L. M., Reactions inside Nanoscale Protein Cages. *Nanoscale* **2011,** *3*, 2376-2389.

13. King, N. P.; Lai, Y.-T., Practical Approaches to Designing Novel Protein Assemblies. *Curr. Opin. Struct. Biol.* **2013,** *23*, 632-638.

14. Comellas-Aragonès, M.; Engelkamp, H.; Claessen, V. I.; Sommerdijk, N. A. J. M.; Rowan, A. E.; Christianen, P. C. M.; Maan, J. C.; Verduin, B. J. M.; Cornelissen, J. J. L. M.; Nolte, R. J. M., A Virus-Based Single-Enzyme Nanoreactor. *Nat. Nanotechnol.* **2007,** *2*, 635-639.

15. Patterson, D. P.; Prevelige, P. E.; Douglas, T., Nanoreactors by Programmed Enzyme Encapsulation inside the Capsid of the Bacteriophage P22. *ACS nano* **2012,** *6*, 5000-5009.

16. Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L. P.; Pedersen, J.; Birkedal, V.; Besenbacher, F.; Gothelf, K. V.; Kjems, J., Self-Assembly of a Nanoscale DNA Box with a Controllable Lid. *Nature* **2009,** *459*, 73-76.

17. Zhao, Z.; Fu, J.; Dhakal, S.; Johnson-Buck, A.; Liu, M.; Zhang, T.; Woodbury, N. W.; Liu, Y.; Walter, N. G.; Yan, H., Nanocaged Enzymes with Enhanced Catalytic Activity and Increased Stability against Protease Digestion. *Nat. Commun.* **2016,** *7*, 10619.

18. Linko, V.; Eerikäinen, M.; Kostiainen, M. A., A Modular DNA Origami-Based Enzyme Cascade Nanoreactor. *Chem. Commun.* **2015,** *51*, 5351-5354.

19. Zhu, Y.; Wang, F.; Zhang, C.; Du, J., Preparation and Mechanism Insight of Nuclear Envelope-Like Polymer Vesicles for Facile Loading of Biomacromolecules and Enhanced Biocatalytic Activity. *ACS Nano* **2014,** *8*, 6644-6654.

20. Wörsdörfer, B.; Woycechowsky, K. J.; Hilvert, D., Directed Evolution of a Protein Container. *Science* **2011,** *331*, 589-592.

21. Hickman, S. J.; Ross, J. F.; Paci, E., Prediction of Stability Changes Upon Mutation in an Icosahedral Capsid. *Proteins: Struct. Funct. Bioinf.* **2015,** *83*, 1733-1741.

22. Padilla, J. E.; Colovos, C.; Yeates, T. O., Nanohedra: Using Symmetry to Design Self Assembling Protein Cages, Layers, Crystals, and Filaments. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** *98*, 2217-2221.

23. King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D., Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science* **2012,** *336*, 1171-1174.

24. Lai, Y.-T.; Cascio, D.; Yeates, T. O., Structure of a 16-Nm Cage Designed by Using Protein Oligomers. *Science* **2012,** *336*, 1129-1129.

25. Lai, Y.-T.; Reading, E.; Hura, G. L.; Tsai, K.-L.; Laganowsky, A.; Asturias, F. J.; Tainer, J. A.; Robinson, C. V.; Yeates, T. O., Structure of a Designed Protein Cage That Self-Assembles into a Highly Porous Cube. *Nat. Chem.* **2014,** *6*, 1065-1071.

26. Bale, J. B.; Gonen, S.; Liu, Y.; Sheffler, W.; Ellis, D.; Thomas, C.; Cascio, D.; Yeates, T. O.; Gonen, T.; King, N. P., Accurate Design of Megadalton-Scale Two-Component Icosahedral Protein Complexes. *Science* **2016,** *353*, 389-394.

27. Sciore, A.; Su, M.; Koldewey, P.; Eschweiler, J. D.; Diffley, K. A.; Linhares, B. M.; Ruotolo, B. T.; Bardwell, J. C.; Skiniotis, G.; Marsh, E. N. G., Flexible, Symmetry-Directed Approach to Assembling Protein Cages. *Proc. Natl. Acad. Sci. U.S.A.* **2016,** *113*, 8681-8686.

28. Raman, S.; Machaidze, G.; Lustig, A.; Aebi, U.; Burkhard, P., Structure-Based Design of Peptides That Self-Assemble into Regular Polyhedral Nanoparticles. *Nanomed. Nanotechnol. Biol. Med.* **2006,** *2*, 95-102.

29. Fletcher, J. M.; Harniman, R. L.; Barnes, F. R. H.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N., Self-Assembling Cages from Coiled-Coil Peptide Modules. *Science* **2013,** *340*, 595-599.

30. Castelletto, V.; de Santis, E.; Alkassem, H.; Lamarre, B.; Noble, J. E.; Ray, S.; Bella, A.; Burns, J. R.; Hoogenboom, B. W.; Ryadnov, M. G., Structurally Plastic Peptide Capsules for Synthetic Antimicrobial Viruses. *Chem. Sci.* **2015,** *7*, 1707-1711.

31. Hsia, Y.; Bale, J. B.; Gonen, S.; Shi, D.; Sheffler, W.; Fong, K. K.; Nattermann, U.; Xu, C.; Huang, P.-S.; Ravichandran, R., Design of a Hyperstable 60-Subunit Protein Icosahedron. *Nature* **2016,** *535*, 136-139.

32. King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D., Accurate Design of Co-Assembling Multi-Component Protein Nanomaterials. *Nature* **2014,** *510*, 103-108.

33. Fletcher, J. M.; Boyle, A. L.; Bruning, M.; Bartlett, G. J.; Vincent, T. L.; Zaccai, N. R.; Armstrong, C. T.; Bromley, E. H. C.; Booth, P. J.; Brady, L. R.; Thomson, A. R.; Woolfson, D. N., A Basis Set of De Novo Coiled-Coil Peptide Oligomers for Rational Protein Design and Synthetic Biology. *ACS Synth. Biol.* **2012,** *1*, 240-250.

34. Thomas, F.; Boyle, A. L.; Burton, A. J.; Woolfson, D. N., A Set of De Novo Designed Parallel Heterodimeric Coiled Coils with Quantified Dissociation Constants in the Micromolar to Sub-Nanomolar Regime. *J. Am. Chem. Soc.* **2013,** *135*, 5161-5166.

35. Suzuki, T.; Arai, S.; Takeuchi, M.; Sakurai, C.; Ebana, H.; Higashi, T.; Hashimoto, H.; Hatsuzawa, K.; Wada, I., Development of Cysteine-Free Fluorescent Proteins for the Oxidative Environment. *PLoS ONE* **2012,** *7*, e37551.

36. Hodgson, L.; Tavaré, J.; Verkade, P., Development of a Quantitative Correlative Light Electron Microscopy Technique to Study Glut4 Trafficking. *Protoplasma* **2014,** *251*, 403-416.

37. Goepfert, P. A.; Tomaras, G. D.; Horton, H.; Montefiori, D.; Ferrari, G.; Deers, M.; Voss, G.; Koutsoukos, M.; Pedneault, L.; Vandepapeliere, P., Durable Hiv-1 Antibody and T-Cell Responses Elicited by an Adjuvanted Multi-Protein Recombinant Vaccine in Uninfected Human Volunteers. *Vaccine* **2007,** *25*, 510-518.

38. Schwarze, S. R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S. F., *In Vivo* Protein Transduction: Delivery of a Biologically Active Protein into the Mouse. *Science* **1999,** *285*, 1569-1572.

39. Minten, I. J.; Hendriks, L. J.; Nolte, R. J.; Cornelissen, J. J., Controlled Encapsulation of Multiple Proteins in Virus Capsids. *J. Am. Chem. Soc.* **2009,** *131*, 17771-17773.

40. Pollok, B. A.; Heim, R., Using Gfp in Fret-Based Applications. *Trends Cell Biol.* **1999,** *9*, 57-60.

41. Akrap, N.; Seidel, T.; Barisas, B. G., Förster Distances for Fluorescence Resonant Energy Transfer between Mcherry and Other Visible Fluorescent Proteins. *Anal. Biochem.* **2010,** *402*, 105-106.

42. Hsieh, H.-J.; Liu, P.-C.; Liao, W.-J., Immobilization of Invertase *Via* Carbohydrate Moiety on Chitosan to Enhance Its Thermal Stability. *Biotechnol. Lett.* **2000,** *22*, 1459-1464.

43. Cunha, A. G.; Fernández-Lorente, G.; Bevilaqua, J. V.; Destain, J.; Paiva, L. M.; Freire, D. M.; Fernández-Lafuente, R.; Guisán, J. M., Immobilization of Yarrowia Lipolytica Lipase—a Comparison of Stability of Physical Adsorption and Covalent Attachment Techniques. *Appl. Biochem. Biotechnol.* **2008,** *146*, 49-56.

44. Mateo, C.; Palomo, J. M.; Fernandez-Lorente, G.; Guisan, J. M.; Fernandez-Lafuente, R., Improvement of Enzyme Activity, Stability and Selectivity *Via* Immobilization Techniques. *Enzyme Microb. Technol.* **2007,** *40*, 1451-1463.

45. Torres, R.; Ortiz, C.; Pessela, B. C.; Palomo, J. M.; Mateo, C.; Guisán, J. M.; Fernández-Lafuente, R., Improvement of the Enantioselectivity of Lipase (Fraction B) from Candida Antarctica *Via* Adsorpiton on Polyethylenimine-Agarose under Different Experimental Conditions. *Enzyme Microb. Technol.* **2006,** *39*, 167-171.

46. Avnir, D.; Coradin, T.; Lev, O.; Livage, J., Recent Bio-Applications of Sol–Gel Materials. *J. Mater. Chem.* **2006,** *16*, 1013-1030.

47. Schoevaart, R.; Wolbers, M.; Golubovic, M.; Ottens, M.; Kieboom, A.; Van Rantwijk, F.; Van der Wielen, L.; Sheldon, R., Preparation, Optimization, and Structures of Cross- Linked Enzyme Aggregates (Cleas). *Biotechnol. Bioeng.* **2004,** *87*, 754-762.

48. You, C.; Zhang, Y.-H. P., Self-Assembly of Synthetic Metabolons through Synthetic Protein Scaffolds: One-Step Purification, Co-Immobilization, and Substrate Channeling. *ACS Synth. Biol.* **2012,** *2*, 102-110.

49. Zhang, Y.-H. P., Substrate Channeling and Enzyme Complexes for Biotechnological Applications. *Biotechnol. Adv.* **2011,** *29*, 715-725.

50. O'Neil, A.; Prevelige, P. E.; Douglas, T., Stabilizing Viral Nano-Reactors for Nerve-Agent Degradation. *Biomater. Sci* **2013,** *1*, 881-886.

51. Villalobos, A.; Ness, J. E.; Gustafsson, C.; Minshull, J.; Govindarajan, S., Gene Designer: A Synthetic Biology Tool for Constructing Artificial DNA Segments. *BMC Bioinf.* **2006,** *7*, 285.

52. Hall, T., Bioedit: An Important Software for Molecular Biology. *GERF Bull Biosci* **2011,** *2*, 60-1.

53. Benchling [Biology Software]: Retrieved from http://benchling.com., 2017.

54. Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M., Selective Reduction of Disulfides by Tris (2-Carboxyethyl) Phosphine. *J. Org. Chem* **1991,** *56*, 2648-2650.

55. Laue, T.; Shah, B.; Ridgeway, T.; Pelletier, S., Computeraided Interpretation of Analytical Sedimentation Data for Proteins.(1992) Analytical Ultrancentrifugation in Biochemistry and Polymer Science. *Roy Soc Chem*, 90-125.

56. Brown, P. H.; Schuck, P., Macromolecular Size-and-Shape Distributions by Sedimentation Velocity Analytical Ultracentrifugation. *Biophys. J.* **2006,** *90*, 4651-4661.

57. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B., Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat. Methods* **2012,** *9*, 676-682.

58. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W., Nih Image to Imagej: 25 Years of Image Analysis. *Nat. Methods* **2012,** *9*, 671.

59. Warren, S. C.; Margineanu, A.; Alibhai, D.; Kelly, D. J.; Talbot, C.; Alexandrov, Y.; Munro, I.; Katan, M.; Dunsby, C.; French, P. M., Rapid Global Fitting of Large Fluorescence Lifetime Imaging Microscopy Datasets. *PLoS ONE* **2013,** *8*, e70687.

For Table of Contents Only