SUPPORTING INFORMATION

Digital Switching of Local Arginine Density in a Genetically Encoded Self-Assembled Polypeptide Nanoparticle Controls Cellular Uptake

Sarah R. MacEwan[1] and Ashutosh Chilkoti[1,2,]*

[1]Department of Biomedical Engineering, Duke University, Box 90281, Durham, NC 27708

[2]Center for Biologically Inspired Materials and Material Systems, Duke University, Box 90271

Durham, NC 27708

*To whom correspondence should be addressed. Email: chilkoti@duke.edu

Materials. All oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). Restriction enzymes, calf intestinal alkaline phosphatase (CIP), and molecular biology buffers were obtained from New England BioLabs (Ipswich, MA). T4 DNA Ligase was purchased from Invitrogen (Grand Island, NY). Competent EB5α and BL21 *E. coli* cells were obtained from EdgeBio (Gaithersburg, MD). Terrific Broth (TB) dry media was purchased from MO BIO Laboratories (Carlsbad, CA) and kanamycin antibiotic was purchased from EMD Chemicals (Rockland, MA). Alexa Fluor 488 C5-maleimide, Alexa Fluor 594 wheat agglutinin, and Hoechst 33342 were obtained from Invitrogen (Grand Island, NY). TCEP-HCL was purchased from Thermo Scientific (Rockford, IL). Dansylcadaverine, genistein, and 5-(N-ethyl-N-isopropyl)amiloride were obtained from Sigma (St. Louis, MO). HeLa and MCF7 cells were purchased from ATCC (Manassas, VA), while primary pooled HUVEC cells, HUVEC media, and supplements were acquired from Lonza (Walkersville, MD). Base media for HeLa and MCF7 cells was purchased from Sigma (St. Louis, MO). All other cell culture reagents were procured from Gibco (Grand Island, NY). Chambered coverglass for live cell confocal microscopy was purchased from Electron Microscopy Sciences (Hatfield, PA).

Iterative design of ELP_{BC}

The challenge in the design of an ELP diblock copolymer (ELP_{BC}) that has a specified critical micelle temperature (CMT) is that, although we can predict the T_t of an individual ELP segment with considerable precision, the prediction of the CMT of an ELP_{BC} is confounded by the interactions between the two blocks. This interaction, as well as a secondary effect that can arise from interactions with the peptide sequence appended to the N- and C-terminus of the ELP, can cause the CMT to differ considerably from the T_t of the hydrophobic block. The design of a suitable ELP_{BC} that displays a CMT of interest hence involves an iterative design cycle that requires systematic adjustment of ELP_{BC} parameters including ELP sequence and block length. Using our previously developed correlation to predict the T_t of an individual ELP segment as a function of its sequence[1] and chain length,[2] we selected a hydrophobic block with a predicted T_t close to the

desired CMT, knowing that the actual CMT of the ELP_{BC} would differ from the T_t as a result of fusion with the hydrophilic block.[3] Next, the sequence of the hydrophilic block was chosen to provide an ELP segment that has a T_t significantly higher than the T_t of the hydrophobic block. The length of the hydrophilic block was restricted such that a 2:3 to 3:2 ratio of hydrophobic:hydrophilic block length was maintained to satisfy this ELP_{BC} parameter that is known to allow micelle assembly.[3] Fusion of the synthetic genes that encode these blocks, and their expression in *E. coli* from a plasmid-borne gene, provided a 'first generation' ELP_{BC} that revealed the effect of the interaction between blocks on the CMT. In the event that the CMT was not at the desired temperature, as was our experience with the 'first generation' ELP_{BC} designed herein, the thermal properties of the 'first generation' design, along with our ability to predict the T_t of each block individually, provided sufficient insight to empirically design a new ELP_{BC} with a CMT in the desired range between 39 °C and 42 °C in one or two subsequent cycles. Figure S1 shows the iterative design cycle used to optimize the thermal properties of the ELP_{BC}.

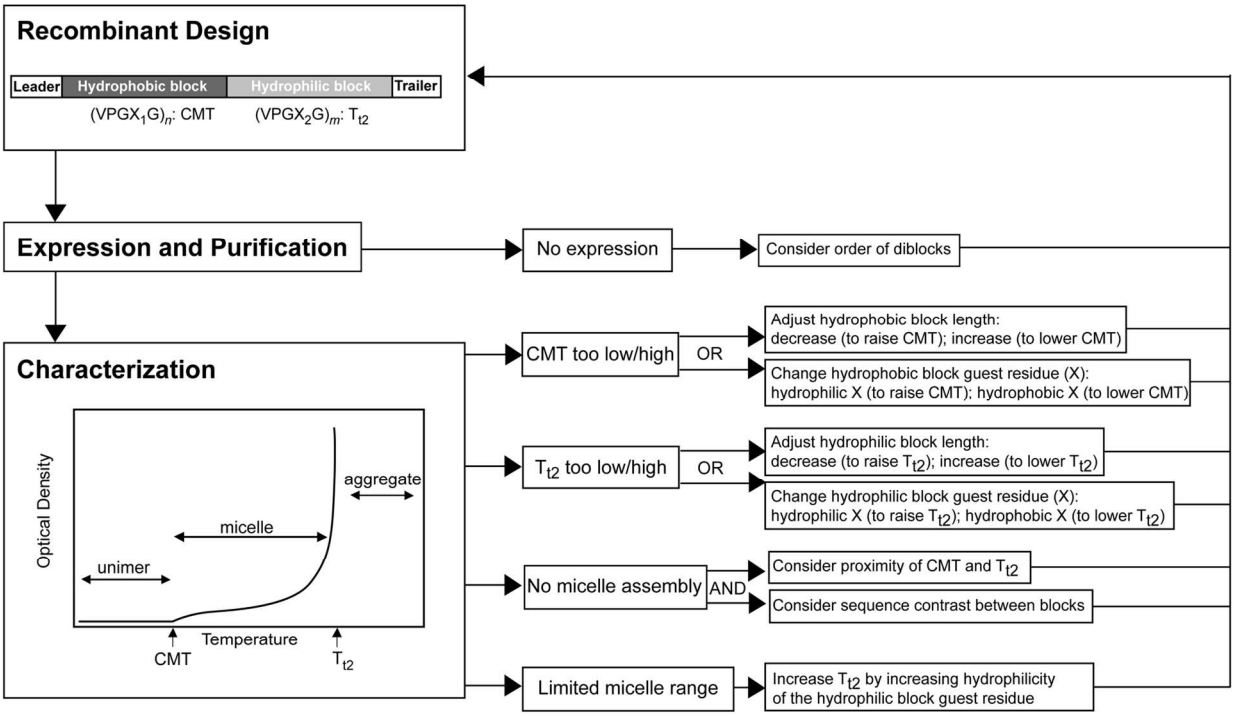

Figure S1. An iterative design cycle was used to optimize the thermal properties of the ELP_{BC}. A 'first generation' ELP_{BC} was recombinantly designed. If the ELP_{BC} was successfully expressed it was purified and characterized with temperature-programmed turbidimetry. Important thermal characteristics of the ELP_{BC}, such as the CMT and T_{t2} (the micelle-to-aggregate transition temperature), were evaluated for the desired application of the ELP_{BC}. In this instance the desired property was a CMT between 39 °C and 42 °C. If the ELP_{BC} did not display the desired thermal properties, adjustments were made to the genetic design using our knowledge of established methods to predictably manipulate the T_t of each individual ELP block. A new ELP_{BC} was recombinantly designed, expressed, and characterized to evaluate the effect of the adjustment on the thermal properties. This process was repeated until an ELP_{BC} with the desired properties was obtained.

Previously designed ELP_{BC}s with a CMT between 39 °C and 42 °C were expressed with a syntax of N-hydrophilic-hydrophobic-C where N and C refer to the polypeptide termini.[3] This orientation required placement of the Arg_5 peptide adjacent to the N-terminus of the ELP. N-terminal functionalization with multiple consecutive Arg residues destroyed expression of ELP in our optimized *E. coli* system. We therefore began our design cycle with an ELP_{BC} that was reversed to place the Arg_5 sequence at the C-terminus. ELP_{BC} genes encoding a polypeptide with a syntax of N-hydrophobic-hydrophilic-C were encoded with Arg_5 peptide at the C-terminus of the ELP. This 'first generation' gene successfully expressed the desired ELP_{BC} in *E. coli*. However, reversing the gene syntax significantly perturbed the thermal properties of the ELP_{BC} such that the CMT was 31 °C, below our desired range for hyperthermia-triggered micelle assembly. The gene design of the 'first generation' ELP_{BC}s was then systematically manipulated to attain a CMT between 39 °C and 42 °C. Parameters manipulated in subsequent design cycles included the ELP block length (as length is inversely proportional to T_t of an ELP) and the ELP sequence (as hydrophilic guest residues (X) in the VPGXG repeat increase the T_t and hydrophobic guest residues lower the T_t). To increase the

CMT to our desired temperature range in the 'second generation' design, the hydrophobic block length was decreased (raising the T_t of the hydrophobic block), resulting in an increase in the CMT to 42 °C, and a decrease in the effective micelle range as the higher CMT was now closer to the T_{t2}, defining the micelle to aggregate transition. The micelle range was extended in the 'third generation' design to promote stability in the self-assembled micelle state. This was achieved by increasing the hydrophilicity of the hydrophilic block sequence by substitution of more hydrophilic guest residues. The length of the hydrophobic block was increased slightly to compensate for this increased hydrophilicity. The resulting 'third generation' design produced an optimized ELP_{BC} with a CMT of 39 °C, suitable for mild hyperthermia-triggered micelle self-assembly.

Recombinant Design of ELP constructs. ELPs were designed directly in a pET24a+ expression vector (Novagen; Madison, WI) by recursive directional ligation by plasmid reconstruction (PRe-RDL) in EB5α *E. coli* competent cells (EdgeBio; Gaithersburg, MD), as described elsewhere.[4] This technique allowed the modular design of ELP_{BC} genes and simple addition of N- and C-terminal sequences that encode leader and trailer peptides, respectively. ELP sequences are defined with the nomenclature $(Val\text{-}Pro\text{-}Gly\text{-}X\text{-}Gly)_n$, where (Val-Pro-Gly-X-Gly) is the repeating pentapeptide unit of the ELP sequence, X is a 'guest' residue that can be any residue other than Pro, and *n* is the number of pentapeptide repeats that determines the overall length of the ELP. The optimized ELP_{BC}s were composed of a hydrophobic $(Val\text{-}Pro\text{-}Gly\text{-}Val\text{-}Gly)_{40}$ segment and a hydrophilic $(Val\text{-}Pro\text{-}Gly\text{-}X\text{-}Gly)_{60}$ segment, in which the guest residue X is Ala or Gly at a ratio of 1:1. The N-terminus of the hydrophobic segment was functionalized with the leader sequence Met-Gly-Cys-Gly-Trp-Pro-Gly, which provides a reactive cysteine residue for attachment of a fluorophore by maleimide conjugation, and a tryptophan residue for quantification of peptide concentration by UV absorbance. The C-terminus of the hydrophilic domain was functionalized with a trailer sequence that included a flexible Pro-Gly-Gly-Ser linker, followed by the Arg_5 sequence to create the Arg_5-ELP_{BC} construct. This Arg_5 oligomer is below the number of Arg residues that are thought to

constitute the threshold to enable efficient cellular uptake.[5] A non-functionalized ELP_{BC} was also synthesized, termed ELP_{BC}, whose hydrophilic domain terminated in the linker sequence described above. This ELP_{BC} served as a negative control that was capable of micelle assembly, but lacked Arg residues. Finally, a size-matched unimer ELP was synthesized with the sequence (Val-Pro-Gly-X-Gly)$_{100}$, in which the guest residue X is Ala or Gly at a ratio of 1:1, equivalent to the ELP sequence utilized for the hydrophilic domain of the ELP_{BC}. This unimer control, termed Arg_5-ELP, served as an Arg-functionalized control incapable of micelle assembly due to its homopolymer design and high T_t.

ELP synthesis and purification. Competent BL21(DE3) *E. coli* cells (EdgeBio; Gaithersburg, MD) transformed with the pET24a+ expression vector bearing the desired ELP gene were grown in TB media (MO BIO Laboratories; Carlsbad, CA) for 24 hours at 37 °C. *E. coli* cells were collected by centrifugation for 10 min at 3,000 RPM, resuspended in phosphate buffered saline (PBS), and lysed by sonication. Polyethylenimine (PEI) was added to the cell lysate to promote precipitation of DNA, and insoluble contaminants were removed by centrifugation for 15 min at 14,000 RPM and 4 °C. ELP was then purified from cell lysate by inverse transition cycling.[6] First, the ELP transition was induced with the addition of up to 3 M NaCl. Coacervated ELP was collected by centrifugation for 15 min at 14,000 at 25-37 °C (Hot spin) and soluble contaminants in the supernatant were discarded. The ELP pellet was resuspended in cold PBS and centrifuged for 15 min at 14,000 RPM at 4 °C (cold spin). The ELP supernatant was retained, removing insoluble contaminants in the pellet. Alternating hot and cold spin cycles were repeated 5 times to achieve pure samples as determined by SDS-PAGE.

Fluorophore conjugation. ELP was fluorescently labeled with Alexa Fluor 488. The reactive cysteine residue at the ELP N-terminus permitted conjugation of fluorophore via maleimide chemistry. Purified ELP at 250 μM was reacted with an excess of 1 mM Alexa Fluor 488 C5-maleimide (Invitrogen; Grand Island, NY) in 10% DMSO, 3 μM TCEP-HCL, and 10 μM $NaPO_4$ at pH 7

for 2 hours at room temperature. ELP was collected from the reaction solution by the addition of NaCl to induce ELP coacervation, and centrifugation for 10 min at 14,000 RPM to collect aggregated ELP. Free fluorophore in the supernatant was discarded and the ELP pellet was resuspended in PBS. Resuspended ELP was passed through a PD-10 desalting column (GE Healthcare; Piscataway, NJ) to remove remaining free fluorophore. Labeling efficiency was determined by absorbance at 488 nm, compensating for fluorophore absorbance occurring at the 280 nm protein peak. Labeled ELP was mixed with respective unlabeled ELP to achieve a labeling efficiency of 50%.

Characterization of temperature-triggered micelle assembly. ELP thermal behavior was characterized by temperature-programmed turbidimetry, dynamic light scattering (DLS), and zeta potential measurements. The optical density of Arg_5-ELP_{BC} and controls was measured at 350 nm from 25 to 60 °C on a Cary 300 Bio UV-Visible spectrophotometer at a concentration of 10 μM in PBS. Prior to DLS, samples were filtered through 0.02 μm filters (Whatman; Piscataway, NJ) at a concentration of 10 μM for fluorophore-labeled ELP or 20 μM for fluorophore-labeled ELP in the presence of endocytosis inhibitors. DLS measurements were performed on a Wyatt Technology DynaPro temperature controlled microsampler over a range of 25 to 55 °C. Zeta potential measurements were obtained on a Malvern Zetasizer Nano ZS90. Due to the need to eliminate buffer salts for zeta potential measurements, the thermal behavior of Arg_5-ELP_{BC} and ELP_{BC} control was re-examined in H_2O, due to the dependence of ELP T_t on the presence of salt. As expected, the CMT and T_{t2} of the ELPs increased in H_2O in the absence of salts. This increase in the CMT determined the appropriate conditions at which to measure the zeta potentials of Arg_5-ELP_{BC} and ELP_{BC} in their self-assembled micelle state. At these determined conditions of 100 μM ELP at 50 °C, Arg_5-ELP_{BC} micelles and ELP_{BC} micelle controls exhibited zeta potentials of 14.3 ± 0.5 mV and -4.8 ± 0.8 mV (average of three replicates ± SEM), respectively. Arg_5-ELP_{BC} exhibited a positive zeta potential, indicative of the high density of Arg on the micelle corona. The negative zeta potential of

the ELP_{BC} was likely due to the presentation of the negatively charged ELP C-terminus on the corona of this non-functionalized control micelle.

Cell culture. HeLa cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 0.1 M non-essential amino acids, 1.0 mM sodium pyruvate, 100 units/ml penicillin, and 100 μg/ml streptomycin. MCF7 cells were grown in Minimum Essential Medium Eagle with Earle's salts, supplemented with 10% FBS, 0.01 mg/ml bovine insulin, 0.1 M non-essential amino acids, 1.0 mM sodium pyruvate, 100 units/ml penicillin, and 100 μg/ml streptomycin. HUVEC cells were cultured in EBM-2 media supplemented with EGM-2MV SingleQuots (Lonza; Walkersville, MD) and 5% FBS. All cells were maintained at 37 °C and 5% CO_2.

Confocal microscopy. HeLa cells were seeded at $2x10^4$ cells per well on an 8 well coverglass slide (Electron Microscopy Sciences; Hatfield, PA). After 24 hours, the media was removed and replaced with serum-free DMEM (SF-DMEM) containing 15 μM of 50% Alexa Fluor 488-labeled ELP. The slides were incubated at 37 °C or 42 °C for one hour after which the media was replaced with Hank's balanced salt solution (HBSS) containing 5 μg/ml Alexa Fluor 594 wheat agglutinin and 2 μM Hoechst 33342 (Invitrogen; Grand Island, NY) to stain the cell membrane and nuclei, respectively. After 10 minutes, cells were washed 3 times with PBS before the addition of HBSS. The cells were imaged immediately on a Leica SP5 confocal microscope, using a 63X water immersion objective.

Flow cytometry. HeLa, MCF7, or HUVEC cells were seeded at $5x10^4$ cells per well in 12-well tissue culture polystyrene plates. After 24 hours the media was removed and replaced with SF-DMEM containing 10 μM of 50% Alexa Fluor 488-labeled ELP. The plates were incubated at 37 °C or 42 °C for one hour after which the cells were washed once with PBS and removed from the culture surface with 0.05% trypsin/EDTA. The cells were washed by centrifugation for 3 min at 1,200 RPM, once with complete media, and twice more with PBS, before immediate analysis on a BD™ LSR II

flow cytometer. Cells were gated on the live population by forward and side scatter and no less than 10,000 cells were evaluated to determine cellular uptake by measure of mean cellular fluorescence, corrected for autofluorescence of untreated cells.

Inhibition of endocytic pathways. To evaluate the contribution of endocytic pathways on the cellular uptake of ELP, HeLa cells were co-incubated with ELP at 20 μM in combination with dansylcadaverine (50 μM), genistein (100 μM), or amiloride (50 μM), which inhibit clathrin-mediated endocytosis, caveolae-mediated endocytosis, or macropinocytosis, respectively. Following one hour of incubation at 42 °C, cells were analyzed by confocal microscopy or flow cytometry, as described previously. The presence of inhibitors did not disturb micelle assembly as measured by DLS, shown in Table S1.

Table S1. DLS of ELP constructs in the presence of endocytosis inhibitors

		ELP Construct (R_H)[1]		
		Arg_5-ELP_{BC}	**ELP_{BC}**	**Arg_5-ELP**
Inhibitor[2]	**No Treatment**	24.7 ± 0.8 nm	20.7 ± 1.3 nm	5.3 ± 0.4 nm
	Dansylcadaverine	23.4 ± 0.5 nm	21.4 ± 0.4 nm	5.7 ± 0.0 nm
	Genistein	22.3 ± 0.6 nm	20.3 ± 0.4 nm	5.3 ± 0.0 nm
	Amiloride	22.7 ± 0.1 nm	21.1 ± 0.4 nm	5.6 ± 0.1 nm

[1] ELP at 20 μM in PBS at 42 °C. Data represents average R_H of 3 replicates ± SEM

[2] Dansylcadaverine, 50 μM; Genistein, 100 μM; Amiloride, 50 μM

REFERENCES

(1) Urry, D. W. *Chem Phys Lett* **2004**, *399*, 177-183.

(2) Meyer, D. E.; Chilkoti, A. *Biomacromolecules* **2004**, *5*, 846-851.

(3) Dreher, M. R.; Simnick, A. J.; Fischer, K.; Smith, R. J.; Patel, A.; Schmidt, M.; Chilkoti, A. *J Am Chem Soc* **2008**, *130*, 687-694.

(4) McDaniel, J. R.; Mackay, J. A.; Quiroz, F. G.; Chilkoti, A. *Biomacromolecules* **2010**, *11*, 944-952.

(5) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. *Proc Natl Acad Sci U S A* **2000**, *97*, 13003-13008.

(6) Meyer, D. E.; Chilkoti, A. *Nat Biotechnol* **1999**, *17*, 1112-1115.